Introducción a la epistemología contemporánea

Jonathan Dancy

Introducción a la epistemología contemporánea

Jonathan Dancy

Traducido por José Luis Prades Celma
Tecnos, Madrid, 1993
ISBN: 8430923004.
ISBN-13: 9788430923007

Título original:
An Introduction to Contemporary Epistemology, 1985

La paginación se corresponde con la edición impresa. Se han eliminado las páginas en blanco.

letra e

Contenido

Prólogo 13
Introducción 15

PARTE I
CONOCIMIENTO
1. Escepticismo 21
2. Conocimiento 39
3. La teoría condicional del conocimiento 54

PARTE II
JUSTIFICACIÓN
4. Fundamentalismo 71
5. El fundamentalismo y las otras mentes 85
6. Teorías empiristas del significado 105
7. Holismo e indeterminación 118
8. Teorías de la coherencia 132
9. Coherencia, justificación y conocimiento 150

PARTE III
FORMAS DE CONOCIMIENTO
10. Teorías de la percepción 167
11. Percepción: la elección de una teoría 185
12. Memoria 210
13. Inducción 225
14. Conocimiento *a priori* 242
15. ¿Es posible la epistemología? 258

Referencias bibliográficas 275
Índice temático 283

A Sarah

PRÓLOGO

Este libro es una versión ampliada de un curso de epistemología que he impartido durante años en la Universidad de Keele. Por supuesto, es el resultado de numerosas conversaciones con estudiantes y colegas. Les doy las gracias a todos ellos, pero debo especial gratitud a David Bakhurst, David McNaughton y Richard Swinburne, que leyeron la totalidad del libro, a veces diversos borradores, y, con suma amabilidad, me ofrecieron comentarios y sugerencias detallados. Mi padre, John Dancy, también revisó con elegancia todo el libro. Agnus Gellatly, Hanjo Glock, Peter Hacker y Cathy McDaniel leyeron algunas partes e hicieron críticas valiosas. Mi gratitud para todos ellos, y para cuatro informadores anónimos comisionados por Basil Blackwell Ltd., por sus intentos de mejorar el texto; también a Charles Swann por su apoyo constante. John Thompson me persuadió de la necesidad de comenzar a escribir, tanto explícitamente como con su ejemplo; considero que también le estoy agradecido.

Pero mi deuda mayor es con mi esposa Sarah, que me proporcionó una fuente inagotable de aliento después de los días malos y un oído generoso tras los buenos. Siempre he sido un tanto escéptico ante las profesiones de gratitud de los autores para con esposo, esposa, amigo o quien fuere, sin el cual no... Nunca más. Tener que vivir con alguien que tiene la mente ocupada casi por completo en un solo asunto es más de lo que podría pedirse razonablemente. Ciertamente, no estaba incluido en el contrato original.

INTRODUCCIÓN

Como indica su título, este libro está concebido como una introducción a los temas más importantes que se discuten hoy en día bajo el rótulo, más bien confuso, de «epistemología» o teoría del conocimiento. La epistemología es el estudio del conocimiento y de la justificación de la creencia. Entre las cuestiones centrales a las que tratan de responder los epistemólogos están: «¿Qué creencias están justificadas y cuáles no?», «Si hay algo que podemos conocer, ¿qué es?», «¿Cuál es la relación entre conocer y tener una creencia verdadera?», «¿Cuál es la relación entre ver y conocer?». Cuestiones como éstas están en el corazón de la epistemología, pero, por supuesto, ésta va más allá y, como en cualquier otra disciplina filosófica, sus fronteras son más bien difusas.

El libro está concebido para estudiantes de los primeros cursos de una licenciatura en filosofía, pero no es preciso que asuste a lectores de otra procedencia; mi padre, por ejemplo, asegura que lo entiende, aunque es posible que no me agradezca la insinuación de que se trata de un lector representativo. Aunque el libro sea introductorio, no veo razón alguna para ocultar mis propias opiniones, que, por lo demás, no son demasiado idiosincrásicas. Creo que es una ventaja para el alumno el tener algo contra lo que reaccionar, sin que ello me haya llevado a empeñarme en alcanzar conclusiones sólidas en cada tema tratado. Algún material del que aquí he utilizado ha aparecido publicado con anterioridad, y estoy agradecido a los editores del *American Philosophical Quarterly* y de *Analysis* por el permiso para la nueva publicación.

El punto de vista que adopto está casi exclusivamente centrado en la tradición angloamericana, aunque cedo el paso a Hegel en el último capítulo. Al final de éste, doy algunas referencias a textos paradigmáticos de la tradición continental, especialmente de la teoría crítica. En la tradición en la que me muevo, hay que elegir entre dos aproximaciones distintas a las cuestiones epistemológicas. La primera, asociada con Descartes, comienza considerando el reto del escepticismo, la pretensión de que el conocimiento es imposible, y confía en que, al

responder al reto, se verá obligada a exponer la naturaleza del conocer, de la que se seguirá la posibilidad del conocimiento. La segunda, que asocio a Grice (1961), mantiene que Descartes distorsionó la reflexión filosófica con la obsesión patológica por un personaje inexistente denominado «escéptico». En lugar de ello, lo que debemos hacer es investigar directamente la naturaleza del conocimiento (si es que el conocimiento tiene una naturaleza) y de la justificación, con la esperanza de que los resultados que alcancemos nos lleven —de uno u otro modo— a una conclusión sobre la naturaleza del conocimiento. Desde este segundo punto de vista, la intención de la buena epistemología no es la refutación del escepticismo. Ni siquiera es preciso que tal refutación sea uno de sus resultados.

Sigo la primera de estas dos aproximaciones, al menos formalmente, ya que mi primer capítulo trata del escepticismo. A lo largo del libro encontraremos referencias, esporádicas y constantes, a una forma de argumento escéptico. Mi razón o mi excusa pueden estar en el hecho de que hay un nuevo renacimiento del interés en el escepticismo. El tema vuelve a ser el foco de atención del último capítulo, en el que cuestiono la posibilidad misma de la empresa en la que me he embarcado en los catorce capítulos precedentes.

El libro se divide en tres partes, la primera sobre el conocimiento, la segunda sobre la justificación y la tercera sobre las formas de conocimiento. Se podrían plantear muchas objeciones, algunas de ellas justificadas, a esta división. Una sería que el escepticismo no se restringe a la duda sobre la posibilidad del conocimiento. Los argumentos escépticos más interesantes, como sugiero en el primer capítulo, incorporan el mismo tipo de duda sobre la posibilidad de la creencia justificada. Por tanto, el capítulo sobre el escepticismo que aparece en la primera parte podría aparecer con el mismo derecho en la segunda. Otra queja podría ser la de que no está claro en absoluto que los temas tratados en la tercera parte, percepción, memoria, inducción, y conocimiento *a priori*, sean *formas* de conocimiento. Se ha mantenido que son *fuentes* del conocimiento. Podría decirse que el conocimiento *se deriva* de lo que es dado en la percepción, que lo que la percepción proporciona no es en sí mismo conocimiento, sino que se convierte en conocimiento tras algunas transformaciones. Por otra parte, la inducción se parece más a una forma de inferencia que a una forma de conocimiento. Muchos autores eluden esas dificultades con capítulos sobre los *objetos* del conocimiento: el pasado, el futuro, el mundo externo, la necesidad, las otras mentes, etc. Pero creo que esta aproximación contiene sus propias dificultades. Una de las más importantes

sería que cada objeto de conocimiento puede ser conocido de muchas formas diferentes. Por ello, yo hablo de formas de conocimiento, y he permitido que los diferentes objetos posibles de conocimiento salgan a la superficie en el momento oportuno. El problema de las otras mentes, por ejemplo, aparece excepcionalmente en el capítulo 5.

Después del capítulo introductorio sobre escepticismo, la primera parte contiene un capítulo sobre la naturaleza del conocimiento y sus relaciones con la creencia verdadera justificada, y otro centrado en una teoría específica. El primero no puede más que rozar la superficie de un área profundamente trabajada en los últimos años. Incluyo el segundo porque creo que esa teoría, aunque reciente, es más que una moda pasajera y contiene algunas intuiciones fundamentales sobre la naturaleza del conocimiento.

En la segunda parte comparo los méritos de dos puntos de vista sobre la naturaleza de la justificación, el fundamentalismo* y el coherentismo. En los capítulos 4 y 5 expongo los defectos de las formas más crudas de fundamentalismo, y en los capítulos 6 y 7 argumento que tenemos más de una razón general para evitar cualquier tipo de fundamentalismo, una razón que se deriva de la teoría del significado. Con ello se plantean, obviamente, ciertas cuestiones sobre prioridades filosóficas; ¿es posible que la teoría del significado sea prioritaria sobre la epistemología y proporcione una suerte de tribunal de apelación para el debate epistemológico? Al margen de ello, estos dos capítulos son quizá más difíciles que los que los rodean, y recomiendo pasarlos por alto en la primera lectura. Si el lector pasa directamente del capítulo 5 al 8, se encontrará con un marcado contraste entre dos teorías diferentes de la justificación, y todo lo que necesitará saber por el momento es que la función asignada a los capítulos intermedios es la de proporcionar una razón contra cualquier forma de fundamentalismo. Los dos últimos capítulos, 14 y 15, dependen, de algún modo, de lo que se dice en el 6 y el 7. Por ello sugiero también que se omitan en la primera lectura, reservando para ésta los capítulos 1-5 y 8-13.

La tercera parte está dominada por dos capítulos sobre percepción, que quizá constituyen el núcleo del libro. Este tema me parece a la vez el más difícil e importante de la epistemología.

* Los términos ingleses «foundationalism» y «foundationalist» se traducen por «fundamentalismo» y «fundamentalista», que se consideran preferibles a los injustificados anglicismos «fundacionalismo» y «fundacionalista». Por supuesto, debe evitarse cualquier connotación que no sea estrictamente epistemológica de los términos castellanos escogidos. *(N. del T.)*

Al final de cada capítulo hay una breve ampliación titulada «Lecturas complementarias», con dos secciones. (Hay también referencias en el texto principal que a veces se recogen al final del capítulo.) La diferencia entre las dos secciones es que, mientras la primera trata de mostrar por dónde comenzar a ampliar los temas principales del capítulo, la segunda contiene referencias a lecturas que proporcionan un trasfondo adecuado, me permite hablar en general sobre el área implicada y ofrece referencias bibliográficas sobre aspectos del capítulo, para que el lector sepa adonde acudir, si lo desea. Los textos mencionados en la primera sección están, más o menos, al mismo nivel de dificultad que el capítulo al que acompañan. Los mencionados en la segunda pueden ser más difíciles. Las referencias siguen el sistema de autor-fecha [por ejemplo, «Grice (1961)»] y se recogen de nuevo en la sección «Referencias bibliográficas», al final del libro.

Por último, se pretende que el índice temático remedie un defecto: el hecho de que puntos de vista, temas y teorías se sucedan en lugares diferentes a lo largo del libro. Por ejemplo, el antirrealismo se discute en el capítulo 1 y vuelve a aparecer en el 9; las aproximaciones antirrealistas a la percepción se discuten en los capítulos 10 y 11, a la memoria en el capítulo 12 y a la inducción en el 13. Todo ello puede ser confundente: hubiera sido mejor el poder tratar cada tema en un único lugar. Si el lector está interesado en saber lo que ha sucedido, o siente curiosidad por saber lo que sucederá, con relación a un tema específico, debe consultar el índice; espero que éste se lo diga.

PARTE I

CONOCIMIENTO

1. ESCEPTICISMO

1.1. ALGUNAS DISTINCIONES

El escepticismo en su forma más interesante depende siempre de un argumento; cuanto mejor es el argumento, más fuerte es la forma de escepticismo que genera. Dado que depende de un argumento, debe poder ser expresado en forma de conclusión. La conclusión escéptica es la de que el conocimiento es imposible. Nadie sabe nada *de hecho*, porque nadie *puede* saber nada.

Se da un contraste entre el escéptico que ofrece un argumento con esta conclusión y otros dos tipos de escéptico. El primero es quien responde a cada aseveración con la cuestión «¿Cómo lo sabes?» y luego, a cualquier contestación se limita a repetir la cuestión hasta que las respuestas se agotan. Esta cuestión repetida es tremendamente eficaz a la hora de dejar a los demás reducidos a un silencio furioso e impotente, pero no hay mucho que aprender de ella hasta que no sepamos lo que esconde. Por supuesto, hay aquí posibilidades interesantes, como, por ejemplo, las proposiciones siguientes:

1. Nadie sabe que *p* a no ser que pueda decir cómo sabe que *p*.
2. No es posible contestar satisfactoriamente a la cuestión «¿Cómo sabes que *p*?» por el mero procedimiento de volver a aseverar que *p*. Se trata de una petición de principio.

El escéptico que plantea la cuestión sin comprometerse con proposiciones de este tipo no presenta ninguna posición filosóficamente interesante. Sin embargo, cuando sí se da ese compromiso volvemos a una forma de escepticismo que depende de un argumento. Es preciso señalar a este respecto que las dos proposiciones anteriores son dudosas. La segunda, por ejemplo, equivale a la aseveración de que no es posible responder a la cuestión «¿Cómo sabes que tienes dolor?» diciendo simplemente «Porque tengo dolor». Quien da esta respuesta está considerando claramente que en algunos casos funciona, y no debemos sentar la cuestión en contra suya por anticipado.

El segundo tipo de escéptico, por el contrario, no nos ofrece tanto un argumento o una pregunta como una mera actitud. Se trata de una persona obstinada cuya pretensión es la de que, aunque la mayoría de la gente se deje convencer por lo que no son más que débiles evidencias, se necesita algo más para que *él* se convezca. Este escéptico pretende tener patrones más altos de evidencia que la mayoría de nosotros; estigmatiza a los otros como seres a los que se engaña o persuade fácilmente. Su posición se convierte en escepticismo propiamente dicho (es decir, en la posición de que el conocimiento es imposible, más bien que algo más infrecuente de lo que pensamos de ordinario) cuando los patrones se colocan tan altos que no pueden satisfacerse. Sin embargo, para ir más allá de la mera extravagancia y tener algún interés filosófico, no puede limitarse a afirmar que los patrones de evidencia más estrictos son los mejores. Debe aportar algún argumento para defender que los patrones ordinarios no son apropiados. Y el argumento debe apelar tanto a nuestros patrones como a los suyos. Aquí acecha el peligro de la incoherencia: ¿es consistente dar un argumento que apele a los patrones ordinarios de evidencia para demostrar que esos mismos patrones no son apropiados?

La conclusión es aquella con la que comenzábamos: las formas interesantes de escepticismo dependen siempre de un argumento. En esta sección consideraremos tres argumentos escépticos lo suficientemente fuertes como para tomarlos en serio. Antes debemos tener en cuenta algunas distinciones entre diferentes tipos de argumento escéptico.

La primera distinción, y la menos importante, se da entre argumentos escépticos locales y globales. El escepticismo local mantiene que hay razones especiales por las que el conocimiento no es posible en un área determinada, aunque pueda ser posible en otra cualquiera. Las áreas predilectas para el escepticismo local son la ética, la religión y el futuro. Podemos saber cómo están las cosas cuando suceden delante de nuestras propias narices, pero no podemos saber si un acto altruista es moralmente bueno, si Dios existe, o si mañana me tomaré un café con leche para desayunar. Es obvio que el escepticismo local espera encontrar su razón de ser en alguna particularidad especial de este tipo de áreas. Sin embargo, mi experiencia es la de que es muy difícil que el escepticismo local se mantenga en mero escepticismo local. Por ejemplo, el escepticismo ético local tiende a expandirse rápidamente y se convierte en escepticismo general sobre lo que no puede observarse, o sobre la posibilidad del conocimiento científico. El problema es encontrar un argumento convincente para el escepticismo local que no tenga consecuencias expansionistas.

No deja de ser una ventaja que el escepticismo local tienda a convertirse en escepticismo general, dado que los argumentos escépticos globales son, por lo general, más convincentes y efectivos que sus contrapartidas locales. Y lo mismo sucede con la segunda distinción que quiero proponer. Algunos argumentos escépticos atacan directamente la noción de conocimiento, pero no tocan algunas otras nociones relacionadas con ésta, como la noción crucial de creencia justificada. Se podría argumentar, por ejemplo, que para saber es preciso estar seguro, pero que nadie puede estar realmente seguro y, por tanto, que nadie puede saber. Dejando a un lado por el momento la fuerza de ese «realmente», todavía podría confiarse en que, incluso concediendo que una condición necesaria para el conocimiento no se cumple y debemos dejar de hablar de conocimiento, no hay ningún problema en seguir hablando de creencias justificadas. Todavía podríamos seleccionar algunas creencias como creencias justificadas, o como más justificadas que otras, y otras como menos justificadas o completamente injustificadas. El argumento escéptico que estamos considerando nos permite referirnos con pleno derecho a la justificación. Podemos tener la impresión de que el argumento, aunque expresa ciertas particularidades del concepto de conocimiento, nos permite arreglárnoslas perfectamente bien, tanto para los propósitos filosóficos como para los prácticos, con la noción de justificación, a la que el argumento no ataca. Por el contrario, una forma más fuerte de argumento constituye una amenaza para ambas nociones al mismo tiempo y pretende que cualquier deficiencia en la noción de conocimiento está presente de igual modo en la de creencia justificada. Nos encontramos con estas formas de argumento y veremos que son siempre más interesantes que sus contrapartidas más débiles. La afirmación de que ninguna de nuestras creencias sobre el futuro está nunca justificada es más importante y más interesante que la afirmación de que, aunque es muy probable que nuestra creencia en que el sol saldrá mañana sea verdadera y esté justificada, nunca puede decirse que sepamos que el sol saldrá mañana.

La tercera distinción es incluso más importante. Podemos separar los argumentos escépticos que, aunque pretendan privarnos del conocimiento (o incluso de la creencia justificada), nos conceden que *comprendemos* las proposiciones cuya verdad se supone que no conocemos de aquellos otros que afirman que la razón por la que no podemos conocer su verdad radica en que no las comprendemos. Un ejemplo obvio sería el contraste entre la sugerencia de que no podríamos tener ninguna evidencia a favor de la proposición de que Dios existe, por

más que comprendemos esa proposición, y la sugerencia de que la proposición nos es incomprensible y *a fortiori* ni podemos saber que es verdadera ni estar justificados para creerla.

Por supuesto, esta distinción sólo tiene sentido si suponemos la posibilidad de comprender una proposición que nunca pudiéramos creer justificadamente o cuya verdad no pudiéramos conocer nunca. La distinción se viene abajo si adoptamos una teoría de la comprensión que una lo que podemos comprender con lo que podemos reconocer como verdadero. En ese caso, todos los argumentos escépticos relevantes pertenecerán al tipo más fuerte; es decir, pretenderán que ni siquiera comprendemos las proposiciones que pretendemos saber.

A primera vista, un argumento para tratar de mostrar que no comprendemos las proposiciones que pretendemos saber debe ser local más bien que general. La razón es que un argumento global pretendería que no comprendemos nada; esto es ridículo, en primer lugar porque es obvio que sí comprendemos algunas cosas y, en segundo lugar, porque es esencial que comprendamos (y de hecho se espera de nosotros que lo hagamos así) el mismo argumento escéptico.

1.2. TRES ARGUMENTOS ESCÉPTICOS

CEREBROS EN CUBETAS

No sabemos que no somos cerebros, flotando en el líquido contenido en una cubeta de laboratorio, conectados con un computador que nos provee de las experiencias que tenemos en cada momento y bajo el control de algún técnico/científico inteligente (o bondadoso, o malévolo, dependiendo de los gustos de cada cual). No lo podemos saber porque, en el caso de que lo fuéramos y si el científico tuviera éxito, nada en nuestra experiencia nos revelaría que lo somos. Por hipótesis, nuestras experiencias serían idénticas a las de algo que no fuera un cerebro en una cubeta. Dado que cada uno de nosotros sólo puede apelar a su propia experiencia, y como la experiencia es idéntica en cualquiera de las dos situaciones alternativas, nada hay que pueda revelar cuál de las situaciones es la que de hecho se da.

¿Es posible que, aunque no sepamos que no somos cerebros en una cubeta, sí sepamos, a pesar de todo, muchas otras cosas que podrían ser más importantes? Por desgracia, parece que, si no sabemos eso, no hay mucho más que podamos saber. Supongamos que alguien pretende saber que está sentado leyendo un libro. Es de suponer que, si sabe

que está sentado leyendo, también debe saber que no es un cerebro en una cubeta y, por tanto (por un simple *modus tollem*), es de suponer que, al no saber que no es un cerebro en una cubeta, no sabe que está sentado leyendo.

El principio en el que descansa este argumento puede ser formalizado como el principio de cierre bajo la implicación conocida:

$$PC^S: [Sap \ \& \ Sa(p \rightarrow q)] \rightarrow Saq$$

Lo que este principio afirma es que, si *a* sabe que *p* y que *p* implica *q*, en ese caso *a* también sabe que *q;* siempre sabemos que son verdad las proposiciones que sabemos que son consecuencias de una proposición que sabemos. (Es una práctica habitual expresar los principios de este tipo en notación lógica, pero en nuestro caso se escribirá la expresión castellana equivalente. Sin embargo, es importante prestar atención a la notación lógica porque ello facilitará las referencias posteriores. PC^S es *un principio de cierre* porque dice que la transición desde algo conocido a algo que se sabe que es una consecuencia de lo conocido no nos lleva más allá del área cerrada del conocimiento.) Así pues, el principio, cuando *a* no sabe que *q* (~S*aq*) y *a* sabe que *p* implica *q* (S*a*(*p* → *q*)) nos permite inferir que *a* no sabe que *p* (~ S*ap*). Por tanto, parece mostrar que en general, cuando no sabemos que no somos cerebros en una cubeta, tampoco podemos conocer ninguna proposición *p* respecto a la que sí sabemos que, si fuera verdadera, no seríamos cerebros en una cubeta. Hay multitud de argumentos ligeramente diferentes. Por ejemplo, la versión cartesiana que considera *q* como la proposición «Estoy soñando». En la primera de las *Meditaciones metafísicas*, Descartes argumenta que, en la medida en que no sabe que no está soñando, no conoce ninguna proposición *p* respecto a la que sí sabe que, en caso de que fuera verdadera, no estaría soñando.

¿Cómo catalogar este argumento haciendo uso de las distinciones de 1.1? No es un argumento completamente global: admite que el conocimiento es posible, es decir, que es posible saber que *p* → *q*, y usa este hecho a modo de palanca. De hecho, el argumento sólo afecta a esas proposiciones cuya verdad significaría que no somos cerebros en cubetas. Por otra parte, se trata de un argumento fuerte en la medida en que pretende atacar la noción de creencia justificada exactamente en las mismas zonas en las que ataca a la noción de conocimiento. Todavía no hemos mostrado este aspecto, pero podemos hacerlo desarrollando una analogía completa del argumento. Todo lo que precisa-

mos es mostrar, en primer lugar, que nuestra creencia de que no somos cerebros en una cubeta no puede justificarse porque nada en nuestra experiencia contaría como evidencia para esa proposición, y, después, utilizar un principio análogo a PC^s:

$$PC^j: [CJap \ \& \ CJa(p \rightarrow q)] \rightarrow CJaq$$

que mantiene que, si *a* está justificado al creer que *p* y que *p* implica *q*, también estará justificada su creencia de que *q*. PC^j parece tan convincente como puede parecerlo PC^c, si no más.

Sin embargo, el argumento no consigue sugerir que no comprendemos la proposición de que estamos sentados leyendo un libro. Todavía se nos permite entenderla, por más que nunca pudiéramos conocerla ni saber si estamos justificados al creerla. El argumento sólo se convierte en un argumento sobre la comprensión si adoptamos además un punto de vista especial sobre esta última, tal y como ya vimos anteriormente.

EL ARGUMENTO DEL ERROR

Todos hemos cometido equivocaciones, incluso en áreas en las que podemos sentirnos completamente confiados: por ejemplo, simples equivocaciones en matemáticas. Y no hay nada en la situación presente a lo que podamos señalar para que nos revele que esta situación no es una de aquellas respecto a las cuales podemos estar equivocados. No podemos decir qué hay en ella que la haga diferente a las situaciones en las que sí hemos cometido errores. Dado que en ese tipo de situaciones no había conocimiento, ¿cómo podemos decir que ahora sí lo hay? No podemos dar ninguna razón por la que la nueva situación es mejor, a este respecto, que una cualquiera de las antiguas en las que estábamos equivocados.

Este argumento descansa en una versión epistemológica de un principio que es bastante habitual en ética (véase Hare, 1963, pp. 7-16), el principio de universalizabilidad. Se dice que es universalizable un juicio respecto a la bondad moral de una acción, en el sentido en que por el mero hecho de hacer tal juicio nos comprometemos a mantener que cualquier acción similar en ciertos aspectos relevantes será también moralmente buena. Cuando se da otra acción que es semejante a la primera en los aspectos relevantes, o bien la catalogamos asimismo como buena o debemos cancelar el juicio previo de que la pri-

mera acción era buena. ¿Cuáles son los aspectos relevantes que hacen que una acción sea suficientemente similar a la primera? Una acción es similar, en ese sentido, a la primera cuando tiene también las propiedades que constituían las razones para el primer juicio. Ser semejante en los aspectos relevantes no es, pues, ser completamente indistinguible. Al menos, nos encontramos con una restricción con relación a las propiedades que deben ser tenidas en cuenta a este respecto, y es la de que han de ser propiedades cuya presencia o ausencia deben ser detectables por la persona que hace el juicio. Las diferencias entre las acciones que esa persona no pueda reconocer no pueden justificar una diferencia en el juicio. El principio de universalizabilidad nos dice, pues, que en ausencia de diferencias detectables debemos hacer de nuevo el mismo juicio. Debe existir algo a lo que podamos señalar para que quede justificada una diferencia de juicio.

Quizá puedan existir propiedades que trasciendan a toda evidencia posible. Lo que quiere decir que siempre sería posible que estuvieran ausentes aun cuando tuviéramos la mejor evidencia posible de su existencia. Quizá la bondad es una de ellas. Ésta es la razón por la que el principio de universalizabilidad tiene tanta importancia en ética. No podemos suponer que una acción es buena y otra no lo es, a menos que podamos seleccionar otra diferencia entre ellas.

El argumento del error expresa las consecuencias en epistemología de este tipo de punto de vista. Supongamos que ayer yo hubiera afirmado que sabía que iba a llover por la tarde basándome en fundamentos normales (previsiones meteorológicas, las formaciones de nubes, etc.), pero que resultara luego que me había equivocado. Cuando hice mi afirmación, el hecho de que no iba a llover trascendía toda evidencia posible, como sucede en todas las afirmaciones sobre el futuro. Lo que quiere decir que si hoy afirmo saber que lloverá por la tarde basándome en los mismos fundamentos que ayer, debo aceptar que ayer por la mañana también sabía que iba a llover por la tarde. Si, por otra parte, abandono mi pretensión de que ayer sabía tal cosa, ya no puedo afirmar que hoy lo sé. La razón estriba en que el único hecho que justificaría la diferencia no me es accesible: los hechos sobre la lluvia vespertina trascienden a toda evidencia matutina. Por lo que mi aceptación de que ayer por la mañana no sabía me impide afirmar que hoy sí lo sé.

Podemos recorrer de nuevo este argumento desde el punto de vista de un observador exterior. Existe la posibilidad de que, aunque *yo mismo* no estuviera justificado al hacer esos juicios diferentes, otra persona lo podría estar al decir que ayer por la mañana yo no sabía y que

hoy sí sé. Podría ser así, por ejemplo, después de que hoy hubiera llovido y los hechos sobre la lluvia ya no trascendieran a todas las evidencias. Aunque yo no pudiera señalar la diferencia relevante en el momento de hacer los juicios, de hecho estaba equivocado ayer y en lo cierto hoy, lo que bastaría para justificar la diferencia en la descripción de un observador exterior al afirmar que ayer no sabía y hoy sí. Sin embargo, es fácil ver qué diría el escéptico a todo esto, independientemente de que sea más o menos extraño suponer que yo no estaría justificado al hacer una afirmación que otro sí podría hacer con plena justificación. Lo que se sugiere es que yo sabía hoy y no sabía ayer, a pesar de que en el momento de hacer las afirmaciones yo no pudiera indicar ninguna diferencia relevante entre los dos días. Esto muestra que, por todo lo que yo sabía, no iba a llover. ¿Cómo podemos aceptar que yo pueda saber que va a llover si, por todo lo que sé, no va a suceder tal cosa?

La conclusión parecer ser que, si reconozco que alguna vez me he equivocado al afirmar que sabía que *p*, ya no puedo saber que *p* nunca a no ser que muestre alguna diferencia relevante entre las dos situaciones. Y nadie puede decir de mí que sé en un caso y no en el otro, porque, por todo lo que sé, estoy equivocado en ambas ocasiones.

Hasta ahora, sin embargo, hemos restringido el argumento de un modo innecesario a casos en los que, de hecho, yo he cometido errores en el pasado. Pero no es necesario ceñirnos a errores reales del pasado. Podemos verlo en un ejemplo moral. Un ejemplo imaginario podría describirse de tal modo que yo tendiera a afirmar que la acción involucrada es buena. Y este juicio debe ser tan universalizable, deberá ser tan vinculante para mis juicios futuros, como si se hubiera tratado de un caso real y no imaginario. De un modo similar, un caso imaginario en el que yo pretendiera saber que *p*, pero en el que *p* fuera falso, debería impedirme la afirmación de que sé que *p* en un caso nuevo que no sea relevantemente diferente (discernible por mí). Así pues, los casos imaginarios son tan efectivos en este argumento como puedan serlo los reales.

Por supuesto, una característica de nuestro primer argumento escéptico era la de que comenzaba con el supuesto imaginario de que alguien era un cerebro dentro de una cubeta, con la particularidad de que nada podía servir para que la persona en cuestión discerniera esa posibilidad de la realidad. De modo que el argumento que acabamos de construir puede considerarse como una compleja defensa de la primera parte del primer argumento escéptico. Parece demostrar que el supuesto imaginario de que alguien es un cerebro en una cubeta al que

se provoca la experiencia de estar leyendo un libro es una manera completamente efectiva de mostrar que nadie sabe que está leyendo un libro. La diferencia entre los dos argumentos parece estar en el camino que toman a partir de ese punto. El primero usa PC^S para mostrar que nadie sabe nada de lo que se sepa que, si fuera verdadero, no seríamos cerebros en el interior de cubetas. El segundo argumenta de un modo más general que, dado que hemos cometido errores, o los cometeríamos en casos similares, ahora no conocemos.

¿Cómo es de fuerte el escepticismo generado por el argumento del error, en caso de que tenga éxito? Si, como se argumentará en 4.2, no existe ninguna área específica en la que no cometamos errores, el argumento del error ha de ser global más bien que local. Sin embargo, no es algo inmediatamente obvio cómo enunciar un argumento semejante contra la noción de creencia justificada. No podemos argumentar directamente que una creencia falsa no puede estar justificada. De modo que, a no ser que podamos afirmar, como hemos hecho con relación al conocimiento, que la creencia de alguien sólo está justificada cuando la persona en cuestión puede diferenciar las situaciones en las que sus creencias son verdaderas de aquellas en las que son falsas, será imposible concluir que sus creencias verdaderas están injustificadas.

No obstante, quizá sea posible apelar a los movimientos iniciales del argumento en favor de que no tenemos ninguna justificación para creer que no somos cerebros en cubetas. En aquel caso afirmábamos que, si nada en nuestra experiencia podía contar como evidencia de que no somos cerebros en una cubeta, no podíamos decir que la creencia en que no lo somos está justificada. La creencia carece de justificación porque no podemos señalar nada que sugiera si somos o no cerebros en cubetas. Del mismo modo, en nuestro caso podemos decir que la creencia carece de justificación porque nada de lo que podemos señalar sugiere que se trata de una situación en la que la creencia es verdadera más bien que una de las situaciones indistinguibles por nosotros —y, por supuesto, raras— en las que es falsa. Si este movimiento es adecuado, nuestro segundo argumento escéptico ataca la noción de creencia justificada, al menos tanto como lo pueda hacer el primer argumento; de hecho, más aún, puesto que es más global.

Debemos conceder, también, que este argumento, tal y como está expuesto, no afecta en modo alguno al problema de la comprensión. Como anteriormente, necesitamos que una teoría de la comprensión vincule la posibilidad de comprender con la de tener creencias justificadas, o la de conocer, si no queremos que la comprensión sobreviva a

la pérdida de la creencia justificada. Una teoría semejante podría afirmar, por ejemplo, que comprender una proposición es ser capaz de distinguir entre las circunstancias en las que estaríamos justificados para creerla y aquellas en las que no. Existen esas teorías, como veremos más adelante.

LA JUSTIFICACIÓN DE LOS ARGUMENTOS A PARTIR DE LA EXPERIENCIA

¿Tenemos conocimiento de sucesos que no hayamos experimentado o que no estemos experimentando en estos momentos? Suponemos normalmente que nuestra experiencia es una guía fiable sobre la naturaleza de las partes del mundo que estamos observando en un momento dado, y que, en casos favorables, nos proporciona conocimiento. Puedo saber lo que hay en el último cajón de mi mesa de trabajo, o lo que desayunaré mañana, por alguna forma de inferencia inductiva a partir de lo que he observado o de lo que estoy observando. David Hume (1711-1776), filósofo e historiador escocés, planteó de un modo particular si eso es o no así (Hume, edición de 1955, capítulo 4.2). Argumentó que no puedo saber que mi diario está en el cajón (cerrado) de mi mesa a menos que tenga una razón para creer que mi experiencia hace que la proposición sea probable; podemos suponer, quizá, que la experiencia relevante sea mi recuerdo de que puse el diario allí hace cinco minutos y el hecho de que no recuerde haber abierto el cajón desde entonces, junto con mi conocimiento general sobre el comportamiento consistente del mundo de experiencia. Pero sólo tengo una razón para creer que la experiencia hace que la proposición sea probable si tengo alguna razón para creer, en general, que los sucesos que no he observado son semejantes a los que he observado. Lo esencial del argumento de Hume es que es imposible tener razón alguna para esta última creencia, dado que —al no haber ninguna contradicción en el supuesto de que es falsa— no es una creencia analítica o necesariamente verdadera. Y no puedo suponer tampoco que la misma experiencia me ha dado razón para creer que lo no–observado se parecerá a lo observado. Cualquier apelación a la experiencia sienta por anticipado el problema que se plantea, al argumentar *desde* la creencia crucial de que nuestra experiencia es una guía fiable, o *desde* la creencia crucial de que lo no–observado se parece a lo observado, y no *en favor* de esas creencias. Por lo que no puedo tener razón alguna para creer que mi experiencia es una guía fiable, y, por ello, no tengo razón

alguna en favor de creencias sobre sucesos que están más allá de la experiencia, de modo que no puede tener conocimiento sobre ellos.

Vale la pena indicar que el argumento de Hume no intenta derivar una conclusión escéptica a partir del hecho de que yo podría estar equivocado (como sí lo hace el primer argumento), ni a partir del hecho de que he cometido equivocaciones (como hace el segundo argumento). En vez de ello, mantiene que nuestra creencia general de que la experiencia es una guía fiable no puede justificarse, dado que todas las posibles justificaciones sentarían por anticipado la cuestión esencial, al suponer que la experiencia puede revelarnos que la misma experiencia es una guía fiable. Hay un defecto obvio en cualquier argumento que parte de la experiencia para tratar de justificar todos los argumentos que parten de la experiencia.

El escepticismo que se origina en el argumento de Hume no es global, dado que sólo trata de nuestro conocimiento de lo no–observado. El argumento ataca claramente tanto la noción de creencia justificada como la de conocimiento, en esa precisa área, al mantener que no podemos extraer de nuestras observaciones razón alguna en favor de lo que no observamos. No afecta a la noción de comprensión; Hume parece aceptar que comprendemos proposiciones sobre lo que no observamos, aunque argumenta por otros motivos que son, en su mayor parte, falsas.

1.3. UN MODO FÁCIL DE ESCAPAR DEL ESCÉPTICO

Ninguno de los tres argumentos mencionados anteriormente es del tipo más fuerte, dado que ninguno ataca la noción de comprensión. Hay argumentos escépticos locales de este tipo más fuerte, como veremos en el capítulo 5 (nuestro conocimiento de las otras mentes). Y al discutir nuestro conocimiento del pasado y el futuro (capítulos 10 y 11) deberemos tener presentes argumentos en favor de que es imposible concebir un suceso como algo que se da fuera del presente, por ejemplo, en el pasado o el futuro. Pero podríamos suponer que es imposible proporcionar un argumento global de este tipo que fuera convincente. No se trata sólo de que es obvio que comprendemos *algo*; más bien sabemos de antemano que sólo se nos puede llevar a aceptar que no comprendemos nada si, tal y como se espera de nosotros, comprendemos el argumento escéptico. Incluso si no comprendemos el argumento, es cierto que comprendemos la conclusión; de modo que la conclusión debe ser falsa.

Esta escapatoria fácil puede utilizarse contra cualquier argumento escéptico global, tanto si ataca sólo al conocimiento como si ataca también a la creencia justificada. Podríamos decir que el escéptico acepta implícitamente que conoce su conclusión de que el conocimiento es imposible, o que pretende que sus premisas justifican su conclusión de que la creencia justificada es imposible. La primera sugerencia parece poco convincente, pero la última es bastante efectiva. ¿Qué utilidad tendría argumentar que la creencia justificada es imposible, dado que, si tuviéramos razón, perderíamos toda razón en favor de la conclusión?

Estas defensas contra el escéptico intentan evitar el examen detallado de los argumentos presentados y se fijan sólo en la conclusión. Lo hacen así en cualquiera de dos formas posibles. O bien discuten el derecho del escéptico a aseverar su conclusión; o sugieren directamente que la conclusión no puede ser verdadera, de modo que no se les puede obligar a considerar ninguna supuesta razón para que la crean. Ya hemos dado ejemplos del primer tipo. Los ejemplos del segundo tipo deben derivarse de la pretensión de que comprender una u otra proposición es comprender bajo qué condiciones es verdadera y bajo qué condiciones es falsa. Si la conclusión del escéptico fuera verdad, no podríamos tener este tipo de conocimiento; de modo que, si la conclusión del escéptico fuera verdad, no podríamos comprenderla. De lo que se sigue que no podemos entender la conclusión sin darnos cuenta de su falsedad. Creo que está claro que este argumento, si tuviera éxito, convertiría el argumento contra el escepticismo global sobre la comprensión en un argumento contra una forma más débil de escepticismo global sobre el conocimiento. Dado que comprendemos lo que está diciendo el escéptico, debemos tener el tipo de conocimiento que se requiere para tal comprensión.

En mi opinión, el escéptico no debe preocuparse en absoluto por este tipo de argumentos. Debe insistir en que no justifican en modo alguno la negativa a considerar los argumentos que él da como lo que en realidad son. Consideremos, en primer lugar, el caso más débil: incluso si el escéptico fuera tan imprudente como para mantener que cualquier aseveración involucra la pretensión de conocer y que él asevera su conclusión de que el conocimiento es imposible, todavía podría mantener su propia posición. Considera que sus premisas son verdaderas y que existe una inferencia válida a una conclusión también verdadera; las premisas podrían incluir la proposición de que en el pasado ha cometido errores. Puede admitir, imprudentemente, que, al usar esa proposición a modo de premisa, está pretendiendo implíci-

tamente que sabe que es verdadera. Pero todavía insiste en que se sigue de ella, y de otras del mismo tipo, que el conocimiento es imposible. Podemos redescribir su argumento del siguiente modo: si sé esto —y esto sería un caso central del tipo de cosas que sé si sé algo—, entonces sé que el conocimiento es imposible, de modo que, si sé algo, entonces no sé nada. Podemos considerar este argumento de dos formas diferentes. O bien es un ejemplo de prueba por *reducción al absurdo*, en la que se supone que algo es verdadero para probar que es falso; o es una forma de exponer una paradoja interna al concepto de conocimiento, ya que el escéptico puede insistir en que, si un concepto tan básico como el de conocimiento puede usarse para llevarnos válidamente desde premisas que son verdaderas a una conclusión falsa o imposible, en ese caso hay algo de incorrecto en ese concepto. Debe haber alguna suerte de tensión interna en el concepto que debe exponerse más bien que permanecer oculta debajo de la alfombra. Por tanto, cualquier respuesta que consistiera simplemente en señalar que la conclusión es falsa o imposible estaría dando la espalda a lo esencial.

1.4. OTRA RÉPLICA

Una réplica habitual a los dos primeros argumentos escépticos es la de decir que no tenemos ningún motivo para preocuparnos de ellos. Dado que se admite, o se sugiere, que en lo que a cada uno de nosotros nos importa, no habría ninguna diferencia entre la hipótesis de que ahora alguien está sentado leyendo y la de que es sólo un cerebro en una cubeta alimentado con las experiencias pertinentes de estar sentado y leyendo, no nos puede importar en realidad qué es lo verdadero y qué lo falso. Nada de genuina importancia puede depender del hecho de que seamos o no cerebros en una cubeta.

Esta respuesta puede adoptar una forma más fuerte o una más débil, pero tiene algún sentido bajo cualquiera de ellas. Mantiene que lo que para el escéptico es la fuerza de sus argumentos no es en realidad sino su punto más débil. El escéptico insiste en que hay una diferencia entre las dos hipótesis, pero se trata de una diferencia que trasciende a toda evidencia posible, es decir, una diferencia que no podemos distinguir en la práctica. De ello concluye que no podemos saber en cuál de esas situaciones estamos en realidad. La réplica admite que la diferencia trasciende a toda evidencia posible, y usa ese hecho en contra del escéptico. Pero ese hecho puede utilizarse de dos formas distintas.

La forma más débil consiste en reconocer que la existencia de una diferencia radical y obvia entre las dos hipótesis es incapaz de establecer diferencia alguna para ninguno de nosotros, de modo que no se nos puede reprochar el que no le prestemos ninguna atención. Es una actitud semejante a la que se podría adoptar ante el problema filosófico de la libertad de la voluntad. Podría intentarse rechazar toda la discusión sobre ese problema, sobre la base de que el hecho de que tengamos o no una voluntad libre no afecta en absoluto al modo en que vivimos nuestras vidas. Actuamos y continuaremos actuando como si tuviéramos una voluntad libre, tanto si nuestras acciones están determinadas como si no lo están. Mi opinión es la de que hay algo profundamente equivocado en esta estrategia, tanto en el área de la libertad de la voluntad como en la del escepticismo. Sin embargo, no es necesario que nos detengamos a examinar sus defectos, dado que, afortunadamente, existe una estrategia más potente y más interesante con la que la anterior puede confundirse.

El escéptico dice que hay una diferencia entre las dos hipótesis, aunque no podamos distinguirla en la práctica porque tal diferencia trasciende a toda evidencia posible. La estrategia débil acepta esto, pero la fuerte lo niega. La posición más fuerte niega que existan verdades o hipótesis que trasciendan a toda evidencia posible, de modo que niega al escéptico el contraste que necesita entre las dos hipótesis. Si la diferencia que se sugiere es una que *no podría* establecer ninguna diferencia para nosotros, se trata de una diferencia completamente vacía y, por tanto, inexistente.

La estrategia débil, entonces, dice que hay una diferencia que no tiene importancia. La fuerte, que no hay ninguna diferencia en absoluto. Podríamos decir que la primera de esas posiciones es *realista*; el realista cree que hay verdades que trascienden a toda evidencia posible, verdades que están más allá de nuestras capacidades efectivas de reconocimiento. La estrategia más fuerte podría denominarse *antirrealista*; niega la existencia de verdades más allá de toda evidencia y sostiene que las diferencias que en un principio no pueden ser reconocidas no existen.

Este tipo de antirrealismo no aparece porque sí, ni es su propósito inicial el de elaborar un contraataque al escepticismo. Pero su espíritu es obvio. Tanto el realista como el escéptico piensan en el mundo como en algo con respecto a lo que, en el mejor de los casos, sólo podemos tener una aprehensión imperfecta. Hay muchos hechos sobre el pasado y el futuro remotos para cuyo reconocimiento o verificación no tenemos medios adecuados. Y siempre es posible que, sin que nos

demos cuenta, el mundo sea completamente diferente a cómo nos parece que es. El antirrealista no cree en la existencia de ese mundo «real» ulterior que estaría por detrás del mundo que sí podemos conocer y que podría ser completamente distinto a nuestro mundo en modos que, si se dieran, no podríamos reconocer en absoluto. Para él, nuestro mundo, el mundo que podemos reconocer, es el único mundo. Es por ello por lo que, para el antirrealista, la tarea de la epistemología es mucho más fácil, ya que se sitúa a los objetos de conocimiento muy cerca de nosotros y, al desaparecer las propiedades trascendentes a toda evidencia, elimina el abismo entre evidencia y verdad. De modo que la cuestión de si en un momento determinado se da o no una propiedad no es distinta a la cuestión de si tenemos la mejor evidencia a favor de su presencia.

El antirrealismo es una teoría de la comprensión que se ha mencionado varias veces a lo largo del presente capítulo. (Tanto el rótulo como su desarrollo en los últimos tiempos se deben a Michael Dummett.) El antirrealista sostiene que la comprensión de las oraciones de nuestro lenguaje debe haberse adquirido en las situaciones que aprendimos a considerar como justificatorias del uso de tales expresiones, situaciones en las que esas expresiones deben contar como verdaderas. Se sigue de ello que, si no existe tal cosa como la creencia justificada, tampoco existe la comprensión. Ya que comprender una oración es ser capaz de seleccionar situaciones que justifican la creencia de que la oración ha de ser verdadera.

Por todo esto, podría parecer que el antirrealismo es una posición particularmente débil. Todo argumento escéptico contra la posibilidad de la creencia justificada es un argumento de la forma más fuerte y nos priva incluso de la comprensión que creíamos poseer respecto a nuestro propio lenguaje. Pero, de hecho, sucede exactamente lo contrario. Los argumentos escépticos que tendrían este efecto requieren un paso que el antirrealista se niega a dar, por lo que nunca alcanzan su devastadora conclusión. Todos ellos requieren la idea de que, sin que pudiéramos darnos cuenta, el mundo podría ser radicalmente diferente a como parece ser y concluyen de ella que no podemos saber que el mundo sea realmente tal y como nos parece que es. Es esto lo que el antirrealista considera imposible. Para él, el sentido de una oración está determinado por los tipos de situación que cuentan a favor la verdad de la oración, de tal modo que la oración con ese sentido (es decir, tal y como la entendemos nosotros) no puede ser falsa cuando se da el tipo de situación que nosotros consideramos que cuenta a favor de su verdad. Así pues, el antirrealismo ofrece una

perspectiva desde la que no sólo no hay ninguna posibilidad de un escepticismo global sobre la comprensión, sino que (por la misma razón) tampoco queda lugar alguno para el escepticismo global sobre la creencia justificada.

El problema en este punto es que el antirrealismo parece ser tan inverosímil como pueda serlo el escepticismo. Para percatarnos de ello, sólo necesitamos comprobar cuáles son las exigencias del antirrealismo y cuan fuertes son nuestras intuiciones realistas.

Un área que parece exigir un punto de vista realista es la de las otras mentes, que discutiremos en el capítulo 5. Nuestra intuición realista en este punto es la de que las sensaciones y pensamientos de los demás son algo que nos está oculto. Por supuesto que observamos su conducta, pero la cuestión de si están experimentando las sensaciones que suponemos que experimentan trasciende, para cada uno de nosotros, cualquier evidencia posible. Siempre es posible que, a pesar de toda su conducta, no se dé, de hecho, ninguna sensación, o que se den sensaciones distintas a las que imaginamos que se dan. De modo que hay un problema real respecto a si hay o no sensaciones en los demás, pero se trata de una cuestión que trasciende a toda evidencia posible.

Otra zona problemática es el pasado. Independientemente de cómo concibamos el futuro, concebimos el pasado como algo que, a su debido momento, ha sido tan determinado como pueda serlo el presente ahora. Pero también asumimos que hay muchas proposiciones sobre el pasado cuya verdad trasciende a toda evidencia posible. A pesar de que no podemos decidir tales verdades, consideramos que descansan sobre hechos definitivos que trascienden a toda evidencia posible y que están ahora más allá de nuestras posibilidades de reconocimiento. Esta actitud respecto al pasado es una actitud realista.

Consideraremos más en detalle estas dos áreas en capítulos posteriores. El objetivo de la discusión precedente ha sido mostrar que, aunque fuera posible, no sería fácil refutar el escepticismo construyendo una alternativa antirrealista al realismo explícito en la actitud escéptica. Si existen áreas en las que el antirrealismo es relativamente aceptable, magnífico. Pero, en la medida en que subsistan otras en las que el realismo parezca ser atractivo, el reto del escéptico mantendrá toda su fuerza. Es posible que no podamos comprar una escapatoria general al escepticismo en el mercado del antirrealismo: el precio que deberíamos pagar sería demasiado alto.

1.5. UNA RESPUESTA MEJOR

Si no existen las escapatorias fáciles, no tenemos más remedio que involucrarnos plenamente en los argumentos que se nos presentan. ¿Dónde podríamos buscar ayudas para una ofensiva crítica? Podría hallarse alguna esperanza en la perspectiva de un análisis satisfactorio del conocimiento que tuviera el efecto de desvelar errores en el razonamiento del escéptico. En el capítulo 3 se considerará un análisis que pretende tener esta consecuencia.

Mientras tanto, quizá debiéramos reflexionar sobre la situación en la que nos quedaríamos si concediéramos que el argumento escéptico es efectivo. Se ha sugerido muchas veces que la epistemología podría prescindir por completo del concepto de conocimiento, porque todas las cuestiones epistemológicas relevantes pueden parafrasearse usando en su lugar el concepto de creencia justificada. Así, en vez de preguntarnos si podemos saber qué sucederá en el futuro, preguntémonos cuándo están justificadas, si es que lo están alguna vez, nuestras creencias sobre el futuro. Y el problema de las otras mentes (capítulo 5) puede exponerse como el problema de cómo mi observación de la conducta de los cuerpos humanos justifica mi creencia de que tales cuerpos son genuinos seres humanos. Poco se gana con la pregunta adicional de si, o cómo, *sé* que son seres humanos.

La dificultad con esta sugerencia es la de que todos los argumentos escépticos presentados, y en realidad cualquier argumento escéptico interesante, parecen dirigirse tanto contra la noción de creencia justificada como contra la de conocimiento. Lo que quiere decir que no se puede mantener esa posición de compromiso fácil. Parece realmente importante para la epistemología que encontremos alguna manera de responder al escéptico.

LECTURAS COMPLEMENTARIAS

El argumento de Descartes sobre el sueño se encuentra en la primera de sus *Meditaciones metafísicas*.

Las reflexiones de Hume sobre la inducción pueden encontrarse en su *Inquiry Concerning Human Understanding*, sección 4, parte 2.

Stroud (1984, capítulo 1) ofrece un análisis sencillo del escepticismo de Descartes.

El argumento sobre cerebros en una cubeta se ofrece en Nozick (1981, pp. 167-171). Sus puntos de vista se analizan más adelante en el capítulo 3.

Stroud (1968) ofrece un valioso análisis de la estrategia antiescéptica considerada

en 1.3, denominada habitualmente trascendental. Lo vincula con el verificacionismo (el «principio de verificación»), una forma anterior de antirrealismo.

Dummett, el arquitecto del antirrealismo como heredero del verificacionismo, ofrece la explicación más introductoria que conozco en Dummett (1978, capítulo 10), pero a los neófitos en este tipo de problemas les recomiendo la lectura del capítulo 9 del presente libro antes de afrontar ese texto.

2. CONOCIMIENTO

2.1. EL ANÁLISIS TRADICIONAL

El análisis tradicional del conocimiento, alrededor del cual gira todo el trabajo reciente al respecto, define el conocimiento como creencia verdadera justificada. Por tanto, mantiene que *a* sabe que *p* si y sólo si

1. *p*,
2. *a* cree que *p*,
3. la creencia de *a* de que *p* está justificada.

Se la denomina la definición tripartita, o el análisis tripartito, por tener tres partes. Define el conocimiento preposicional, o conocimiento de que *p*. No define el conocimiento por familiaridad directa como en «José conoce a Pedro» ni el denominado «saber–cómo», por ejemplo el saber montar en bicicleta, a no ser que pueda mostrarse que son reductibles al conocimiento proposicional.

La definición tripartita tiene atractivos obvios. Normalmente, se considera que la primera cláusula, la que dice que, si *a* sabe que *p*, *p* es verdadera (que puede leerse como S*ap* → *p*), es estipulativa. La segunda cláusula, que, si *a* sabe que *p*, *a* cree que *p* (S*ap* → C*ap*), parece ser de contenido mínimo. La tercera, que, si *a* sabe que *p*, su creencia de que *p* está justificada (S*ap* → CJ*ap*), trata de evitar que cuente como conocimiento cualquier acierto casual por parte de un imprudente que confíe acriticamente en sus propias decisiones. Sin embargo, es importante que tengamos en cuenta una consecuencia de esta justificación de la cláusula 3: normalmente una creencia no está justificada por el mero hecho de ser verdadera, de otro modo la cláusula 3 sería redundante. Si decido la inversión en Bolsa más rentable por el método de echar una moneda al aire y resulta que acierto, el feliz resultado no puede justificar mi elección: no tenía ninguna justificación real para hacer la elección que hice. (También podríamos distinguir entre dos formas de justificación: justificación antes del suceso y

después de él. Utilizaríamos, en ese caso, la primera de ellas en nuestra definición tripartita. Pero, entonces, la cuestión sería si se trata realmente de formas distintas de lo mismo.)

¿Qué problemas plantea la definición tripartita? Podríamos pensar que la cláusula 2 es insuficiente: creer que *p* no es tan fuerte como estar seguro de que *p*, y para saber no basta con creer, uno debe estar seguro.

El mejor motivo para introducir la certeza en nuestro análisis del conocimiento radica en que, muy a menudo, tenemos razón al vacilar respecto a nuestras pretensiones de conocimiento cuando no estamos seguros. Este movimiento de vacilación parece deberse a algo intrínseco al conocimiento, y no habría manera obvia de explicarlo si el conocimiento fuera lo que afirma la definición tripartita. A pesar de que se sugiera muchas veces que la noción de certeza es relevante para el análisis de las *declaraciones* de conocimiento, pero no para el análisis del conocimiento mismo [por ejemplo, Woozley (1953)], hay que tener en cuenta que, en tal caso, no tendríamos modo alguno de explicar por qué la certeza es necesaria para poder declarar que sabemos sin serlo para el conocimiento mismo, es decir, para la existencia de aquello que declaramos.

Dado que vamos a descubrir otras razones para rechazar la definición tripartita, no hay motivo para proseguir ahora con este argumento. La moraleja debe ser la de que, para dar un análisis del conocimiento que no incluya el requisito de la certeza, debemos buscar algún lugar apropiado para la noción de certeza. Un análisis que considere que la certeza es un requisito para las declaraciones de conocimiento necesita poder explicar en sus propios términos por qué ha de suceder tal cosa.

¿Por qué no cambiar sin más la cláusula 2 por «a está seguro de que *p*»? Porque estamos dispuestos, en circunstancias que son perfectamente normales, a conceder que alguien sabe algo por más que la persona en cuestión no se atreviera a declarar su conocimiento por carecer de la certeza necesaria. Pensemos en un mal estudiante que, por ejemplo, se ha aprendido la noche anterior las fechas de la muerte de los reyes de Inglaterra durante varios siglos, pero que se siente tan amedrentado por el profesor que pierde toda seguridad sobre la corrección de las respuestas que se le ocurren cuando le pregunta. Supongamos que sus respuestas son las correctas, ¿no diríamos que las sabía, aunque él mismo no se atreviera a decir tal cosa? Parece obvio que nuestras razones para concederle conocimiento están muy cerca de lo que nos sugiere la definición tripartita: da la respuesta correcta, y no lo hace así por mera casualidad.

Hay un problema en este ejemplo del mal estudiante, que está relacionado también con la cláusula 2. En la medida en que el estudiante no está seguro de las respuestas que se le ocurren, ¿diríamos que las cree? Si no vamos con sumo cuidado, este tipo de ejemplos contra el requisito de la certeza puede hacernos perder la condición de la creencia, que tratábamos de salvar.

2.2. LOS CONTRAEJEMPLOS DE GETTIER

Henry está mirando la televisión una tarde de junio. Es la final masculina del torneo de Wimbledon y McEnroe está venciendo a Connors por dos *sets* a uno y con punto de partido en el tercero. McEnroe gana este punto. Henry cree justificadamente que

1. Acabo de ver la victoria de McEnroe en la final de Wimbledon de este año

e infiere razonablemente que

2. McEnroe es el campeón de Wimbledon de este año.

Sin embargo, lo que ha sucedido es que las cámaras instaladas en Wimbledon se han estropeado y lo que muestra la pantalla es una grabación del partido del año anterior. Mientras sucede eso, McEnroe está, de hecho, a punto de repetir la abultada victoria que obtuvo el año pasado. De modo que la creencia 2 de Henry es verdadera, y está ciertamente justificado para tener tal creencia. Pero no concederíamos que Henry sabe 2.

Este tipo de contraejemplo al análisis tripartito del conocimiento es conocido como contraejemplo de Gettier, a partir de E. L. Gettier (1963). Gettier argumentó que estos casos muestran que el análisis tripartito es insuficiente: es posible que alguien no sepa algo, aunque se satisfagan las tres cláusulas pertinentes. Gettier no pone en cuestión ninguna de las tres cláusulas. Concede que individualmente son necesarias, y se limita a argumentar que no son suficientes.

Por motivos que quedarán claros más adelante, es importante que formalicemos la situación. Traduciendo 1 por p, y 2 por q, tenemos:

$$\sim p,\ \mathrm{C}ap,\ \mathrm{CJ}ap,\ p \rightarrow q,\ \mathrm{CJ}a\,(p \rightarrow q),\ q,\ \mathrm{C}aq,\ \mathrm{CJ}aq.$$

De modo que, en un contraejemplo de Gettier, *a* tiene una creencia justificada pero falsa, de la cual infiere la creencia justificada en algo que de hecho es verdad, llegando así a una creencia justificada y verdadera que no es conocimiento.

¿Qué respuesta podemos dar a este tipo de contraejemplos, enormemente conocidos pero todavía un tanto molestos? Parece que se nos abren tres opciones:

1. encontrar algún medio de mostrar que los contraejemplos no funcionan;
2. aceptar los contraejemplos y buscar algún modo de complementar el análisis tripartito de modo que los podamos excluir de él;
3. aceptar los contraejemplos y alterar la adecuación del análisis tripartito, en vez de añadirle algo más.

El resto de esta sección trata de la primera de estas alternativas.

¿En qué principio de inferencia descansan los contraejemplos? El mismo Gettier expone dos. Para que los ejemplos funcionen debe ser posible que una creencia falsa pueda estar justificada; y una creencia justificada debe justificar cualquier creencia que sea implicada por ella (o que se crea justificadamente que es implicada por ella). Esto último es exactamente el Principio de Cierre PC^j, mencionado anteriormente en la discusión del escepticismo (1.2). De modo que, si pudiéramos demostrar que PC^j es falso, nos encontraríamos con el doble efecto de que socavaríamos los contraejemplos de Gettier y (al menos parte de) el primer argumento escéptico. Sin embargo, todavía sería posible construir nuevas variaciones sobre el tema de Gettier que no descansaran en la inferencia, o en inferencias de este tipo, como veremos en la sección siguiente. Si eso es así, ninguna queja sobre PC, ni sobre otros principios, sería muy efectiva.

Lo que no podemos es rechazar los ejemplos de Gettier como si fueran demasiado forzados y artificiales. En sus propios términos, son perfectamente efectivos. Pero sería razonable que nos preguntáramos cuál es el objetivo de exprimirnos el cerebro tratando de encontrar una definición adecuada de «*a* sabe que *p*». ¿Se trata de algo más que de un mero ejercicio técnico? ¿Por qué deberíamos preocuparnos si no pudiéramos dar con una definición libre de todo problema? Muchos de los artículos escritos en respuesta a Gettier dan la impresión de ser parte de un juego entre filósofos que no tiene ningún interés para nadie, excepto los propios jugadores. Después de todo, ¿no

nos ha demostrado Wittgenstein que un concepto puede ser completamente legítimo aunque no tenga ninguna definición, al argumentar que no es necesaria la existencia de un elemento común en todas las instancias de una propiedad (por ejemplo, instancias de conocimiento) distinta a aquella de la que son instancias (por ejemplo, que son conocimiento). [Cf. Wittgenstein (1969b), pp. 17-18, y 1953, §§ 66-67.] De modo que ¿qué demonios importaría que tuviéramos o no éxito en la empresa de descubrir condiciones necesarias y suficientes para el conocimiento?

Simpatizo de muchas maneras con el tono general de esta queja, como resultará evidente muy pronto. Lo que me mantiene en la búsqueda de una respuesta a Gettier es el sentimiento de que puede ser posible encontrar un análisis de lo que sea el conocimiento que tenga un efecto sustancial sobre lo que se dirá más tarde respecto a la justificación. Esto podría suceder en una cualquiera de dos maneras posibles. Podríamos encontrar un análisis de lo que es el conocimiento que fuera suficiente para socavar movimientos esenciales para el escéptico, confirmando con ello la posibilidad de que algunas de nuestras creencias estuvieran justificadas; el análisis que apoyo provisionalmente en el capítulo 3 pretende algo semejante. Otra posibilidad sería la de definir la justificación en términos del conocimiento. Por ejemplo, podríamos suponer que una creencia está justificada si y sólo si en ciertas circunstancias (que deberían ser especificadas) produce conocimiento. (Debo esta idea a Jennifer Hornsby.) Mientras tanto, consideremos algunos análisis del conocimiento que no parecen tan fructíferos.

2.3. RESPUESTAS A GETTIER

Supongo que es obvio que he evitado, en la medida en que he podido, cualquier diagnóstico, por provisional que fuera, sobre el defecto del análisis tripartito descubierto por Gettier. La razón es que todas las respuestas a Gettier surgen de diagnósticos diferentes sobre las carencias del análisis tripartito: una vez que conozcamos lo que falta, será fácil añadirlo.

LA PRESENCIA DE LA FALSEDAD RELEVANTE

El diagnóstico más obvio es fijarse en la falsedad de *p*, que es la creencia inicial de la que se infiere la creencia verdadera y justificada

q. Podríamos, pues, añadir al análisis tripartito la cuarta condición de que no puede ser conocido nada que se haya inferido de una creencia falsa, o de un grupo de creencias entre las que haya una falsa.

Esta sugerencia simple tiene dos defectos. En primer lugar, pueden escribirse variaciones del tema de Gettier en las que, aunque haya falsedad, no se dé ninguna inferencia. Supongamos que creo que hay una oveja en el campo cercano a causa de lo que veo. No estoy infiriendo la existencia de la oveja a partir de lo que veo. Simplemente, veo que hay una oveja. El animal que veo es un perro pastor, pero mi creencia no es falsa porque, oculta por la cerca, hay una oveja que no veo. Podríamos admitir que mi creencia es verdadera y está justificada y rechazar que existiera conocimiento. [El ejemplo proviene de Chisholm(1977),p. 105.]

Una réplica posible sería decir que estoy infiriendo que veo una oveja en el campo a partir de mi conocimiento de mis propios estados sensoriales. Esta respuesta plantea cuestiones de mayor calado, pero en el capítulo 5 se da un argumento bastante claro en favor de que, si hay conocimiento no–inferencial, debe darse a veces sobre cosas distintas a nuestros propios estados sensoriales; por ejemplo, ¿por qué no sobre ovejas?

El segundo defecto radica en que la sugerencia es demasiado fuerte y convierte en improbable que alguien de nosotros sepa algo alguna vez. Como veremos, ése es un peligro común a muchas de las respuestas a Gettier. En este mismo momento, muchos de nosotros tenemos creencias falsas que desempeñan algún papel en nuestras inferencias, de modo que, de acuerdo con la sugerencia, ninguna de nuestras creencias verdaderas y justificadas contaría como conocimiento.

Para eliminar estos defectos, debemos dejar a un lado las referencias a la inferencia y estrechar la relación especificada entre las creencias que son falsas y las que, siendo verdaderas y justificadas, no debieran contar como conocimiento. Por ejemplo, podríamos exigir que no se diera falsedad relevante. Así podríamos afrontar el ejemplo de la oveja en el campo, porque es de presumir que yo crea (con falsedad) que el animal que veo es una oveja, por más que esa creencia no sea usada para ninguna inferencia. Pero esta sugerencia parece identificar la dificultad más bien que resolverla: ¿Qué creencias falsas deben contar como relevantes?

Una respuesta podría ser la de decir que la falsa creencia de que *p* es relevante en el sentido requerido cuando la creencia en *q* habría cesado de estar justificada si el creyente hubiera creído ~*p* en vez de *p*. No todas las creencias falsas son relevantes en este sentido.

Algunas serán tan distantes o tan insignificantes que el hecho diferencial de creer en ellas o en sus opuestas no tendría efecto alguno sobre lo que se cree en un momento determinado. Por ejemplo, entre las creencias por las que creo que Napoleón fue un gran soldado puede haber una que sea falsa, pero tan insignificante que mi justificación para creer que Napoleón fue un gran soldado se mantendría si cambiara de opinión respecto a ella. Una creencia falsa de este tipo no sería relevante en el sentido comentado.

Pero el nuevo análisis plantea dificultades, que pueden ilustrarse mejor por medio de un ejemplo. Supongamos que espero que un colega me lleve a casa en su coche esta tarde, y que el coche tiene la batería descargada. De hecho, no sería un gran problema porque el coche de otro amigo está aparcado al lado con los cables adecuados que podrían usarse para arrancar el coche estropeado. Ahora creo que me llevará esta tarde en su coche y mi creencia está justificada. ¿Sé que me llevará en su coche? El requisito de que no haya creencias falsas relevantes sugiere, plausiblemente, que el que yo sepa o no depende de otras creencias que yo tenga. Pero esta sugerencia plantea dificultades. Si, por ejemplo, me limito a creer

1. me llevará a casa en su coche esta tarde,

se me podría conceder que lo sé, pero si creyera 1 y

2. la batería de su coche no está descargada,

ya no puedo saber 1, dado que tengo una creencia relevante falsa. Sin embargo, si, además de creer 1 y 2, también creo

3. el coche de un amigo está aparcado al lado con cables para cargar baterías,

esta creencia, aparentemente gratuita, tiene el efecto de que yo crea de nuevo que me llevarán a casa en coche. Dado que, si creyera lo contrario de 2, mi creencia 1 no estaría justificada a no ser que creyera también algo similar a 3. Parece, pues, que la sugerencia que analizamos tiene el efecto de que el que yo tenga conocimiento dependerá normalmente de otras creencias aleatorias que yo pueda tener. Hay algo insatisfactorio en este punto, y es necesario que se trabaje más para defender este análisis contra objeciones de este tipo.

REFUTABILIDAD

Un punto de vista ligeramente distinto diagnostica que la fuente de los contraejemplos de Gettier es la presencia de algunas verdades que, de haber sido creídas, habrían destrozado la justificación del creyente [cf. Lehrer y Paxson (1969), Swain (1974)]. Supongamos, por ejemplo, que Henry hubiera creído que estaba mirando una grabación de la final del año anterior, como de hecho sucedía; en ese caso, su justificación para la creencia de que *p* y, por PC^j, para su creencia de que *q* se habría venido abajo. La sugerencia es, pues, que debe añadirse una cuarta cláusula con el requisito de que no puede existir ninguna verdad tal que, si Henry la creyera, perdiera toda justificación para creer en *q*. Ésta es la sugerencia de la refutabilidad; para que haya conocimiento, requeriremos que la justificación no sea refutable, es decir, que no haya verdades cuya adición pudiera refutarla.

Esto no implica que una creencia falsa nunca estará justificada, dado que la sugerencia dice que, aunque algunas creencias están justificadas y sean refutables, lo que necesitamos para que haya conocimiento es una justificación irreprochable. Sin embargo, corremos el peligro de convertir en algo redundante la primera de las condiciones para el conocimiento ($Sap \rightarrow p$). Parece que una creencia falsa nunca podría estar justificada irrefutablemente, dado que siempre habría alguna verdad (como mínimo, la negación de la creencia falsa) cuya adición crearía la justificación. Quizá esto sea una virtud, y no un defecto, dado que el nuevo análisis cuatripartito tiene más coherencia que el anterior; proporciona, en la cuarta cláusula, una explicación de algo que antes se había introducido a modo de mera estipulación: que el conocimiento requiere verdad.

La sugerencia de la refutabilidad puede considerarse como una extensión del requisito anterior de que no hubiera falsedad relevante; estamos mirando, más allá de las proposiciones que el creyente está creyendo, las proposiciones que tendrían determinado efecto en caso de ser creídas. Pero esta ampliación no es ninguna ventaja. Un ejemplo ilustrará mejor el tipo de dificultad con que se enfrenta la noción de refutabilidad. Supongamos que yo crea que mis hijos están ahora jugando en el jardín de casa, y que yo tenga buenas razones para esta creencia. Sin embargo, sin que yo lo sepa, un vecino vino a casa después de que yo saliera por la mañana y les invitó a dar un paseo. Si yo hubiera sabido esto, mi justificación para creer que están en el jardín se vendría abajo, dado que sé también que, normalmente, aceptan este tipo de invitaciones. Pero, en este caso, mi mujer ha rechazado la invi-

tación porque está preocupada por la salud de uno de los niños. Bajo esos supuestos, ¿sé que mis hijos están jugando en el jardín de casa? Si la intuición del lector es afirmativa, deberá rechazar el criterio de refutabilidad tal y corno lo acabamos de formular. Si la respuesta es negativa sobre la base de que, si yo hubiera sabido que el vecino los había invitado, mi justificación habría sido refutada, todavía necesitamos una explicación de por qué la (desconocida para mí) verdad de que mi mujer ha rechazado la invitación no vuelve a equilibrar el fiel de la balanza. En cualquier caso, la sugerencia de la refutabilidad necesita ser modificada.

El problema parece radicar, como en el caso de la exigencia de que no hubiera falsedades relevantes, en la manera en que las nuevas verdades pueden ser añadidas, poco a poco, para destituir la justificación existente, mientras ulteriores verdades todavía esperan en el transfondo para destituir las destituyentes. Lo primero que deberíamos preguntarnos es si no es probable que siempre haya alguna verdad que, si fuera considerada con exclusión de todas las demás, destrozara mi justificación. Incluso si tal cosa no sucede siempre, sucederá lo bastante a menudo como para reducir el alcance de mi conocimiento, lo que, por sí mismo, ya es una objeción. En segundo lugar, necesitamos encontrar una manera de bloquear el modo en que la paulatina adición de verdades nuevas me da y quita conocimiento una y otra vez.

El segundo objetivo podría alcanzarse si alteráramos nuestro análisis de la refutabilidad, de modo que, en vez de hablar de alguna otra verdad (que origina el problema de la adición paulatina de verdades), habláramos de la totalidad de las verdades. Podríamos requerir como cuarta condición que la justificación se mantuviera después de la adición de todas las verdades a nuestro conjunto originario de creencias. Esta nueva noción de refutabilidad parece permitir (probablemente) que yo sepa ahora que mis hijos están jugando en el jardín de casa, porque la segunda verdad añadida elimina los poderes refutantes de la primera. Pero todavía hay problemas en esta nueva noción de refutabilidad. En primer lugar, al hablar de la adición de la totalidad de las verdades parece que ingresamos definitivamente en el terreno de la pura ficción. ¿Tenemos alguna noción adecuada de la «totalidad de las verdades»? En segundo lugar, parece que, según este nuevo criterio, nunca pasaremos de tener razones mínimas para creer que sabemos algo; al creer algo lo que estamos creyendo es que, cuando se añadan todas las verdades, nuestra justificación todavía se mantendrá. A primera vista, no necesitamos tanto para apoyar las pretensiones ordinarias de conocimiento.

FIABILIDAD

Un punto de vista diferente nos aparta de la relación entre la proposición que se pretende saber y otras creencias falsas que deberían ser verdaderas u otras verdades que deberían haber sido creídas. Se sugiere a veces que una creencia justificada puede ser conocimiento cuando se deriva a partir de un método fiable [véanse Goldman (1976); Armstrong (1973), e. 13; Swain (1981)]. En el ejemplo de Gettier, Henry sabe que esta tarde se está jugando la final de Wimbledon. Esta creencia verdadera y justificada se deriva de métodos fiables, como la lectura de los periódicos, que están normalmente en lo cierto con relación a este tipo de cosas. Sin embargo, su creencia de que *q* se deriva de un método que no es fiable. Le hubiera llevado al mayor de los errores si se hubiera dado el caso de que McEnroe hubiera sufrido un inesperado decaimiento, sucumbiendo definitivamente ante los esfuerzos de Connors.

Este punto de vista puede elaborarse mucho más; en algunos extremos, está bastante cerca de los análisis causales que se mencionarán a continuación, porque necesitamos una explicación de qué sea lo fiable, y una respuesta causal es tentadora [véase Goldman (1979)]. Sin embargo, ya podemos ver las dificultades que plantea cualquier variante de este punto de vista. Corre el peligro de o bien hacer imposible el conocimiento, o bien llevarnos directamente a uno de los argumentos escépticos.

Podemos utilizar «fiable» en el sentido de que un método apropiado, cuando se le sigue adecuadamente, es completamente fiable y nunca nos lleva a una creencia falsa. Pero, independientemente de la dificultad general de distinguir entre un defecto en el método y un defecto en la manera en que el método se aplica, parece improbable que haya métodos completamente fiables para la adquisición de creencias. Los seres humanos somos falibles, lo que se muestra no sólo en la manera en que usamos ciertos métodos, sino también en los mismos métodos de adquisición de creencias que nos son accesibles. Podemos decir que, si requiere un método infalible o perfectamente fiable, el conocimiento es imposible.

Si dejáramos a un lado la noción de fiabilidad perfecta y exigiéramos tan sólo que el método fuera fiable en líneas generales, atraeríamos argumentos escépticos del segundo tipo. ¿Cómo es posible que un método que ha fallado una y otra vez en circunstancias relevantemente similares baste en un momento dado para producir conocimiento? Si confiáramos en que el análisis del conocimiento nos ayudara a

rechazar los argumentos escépticos, este particular análisis no nos facilitaría en absoluto las cosas. Nada de esto muestra, por supuesto, que el análisis sea incorrecto. Un análisis correcto del conocimiento podría tener la desagradable consecuencia de darle al escéptico la oportunidad que está buscando. Pero no debemos aceptar que tal cosa sea inevitable hasta que estemos convencidos de que no hay otro análisis del conocimiento que ofrezca menos juego al escéptico. Podemos mantener la esperanza de encontrar uno que le haga la vida más difícil.

Una retirada final sería la de exigir sólo que el método fuera fiable en *este* momento. Esto tiene el efecto de desviar nuestra atención de los casos previos en los que el método ha fallado y evitar así el argumento escéptico que tiene su punto de partida en esos casos. Pero sería razonable dudar de si el requisito de que el método sea fiable en un caso particular equivale a alguna genuina adición al análisis tripartito. Si la fiabilidad es definida en términos de la producción de la verdad, no añade nada a la primera condición, una vez que restringimos nuestra atención al caso particular. Si se la define en términos de justificación, no añade nada a la tercera. Y no parece que haya otras alternativas. (No obstante, bien pudiera ser que la teoría causal equivaliera a una noción de justificación en el caso particular; véase 2.4.)

RAZONES CONCLUYENTES

Otro punto de vista diría que el fallo de Henry en el ejemplo comentado se debe al hecho de que sus razones no llegaban a ser concluyentes. Si, para que haya conocimiento, requerimos que la creencia verdadera justificada esté basada en razones concluyentes, todos los ejemplos de Gettier, y todos los casos en los que el creyente acierta por casualidad, se vienen abajo.

Este punto de vista requiere que se trabaje en un análisis convincente de qué es una razón concluyente. Una sugerencia podría ser la de que cuando las creencias A – M constituyen razones concluyentes para la creencia N, A – M no podrían ser verdaderas si N fuera falsa. Esto excluiría los contraejemplos, pero, en el mejor de los casos, haría del conocimiento un fenómeno extraño. De una u otra forma, el conocimiento empírico sería algo imposible; en el campo de la experiencia, nuestras razones nunca son concluyentes en este sentido.

Un análisis más débil, ofrecido por F. Dretske (1971), sugiere que las razones A – M de alguien en favor de una creencia N son conclu-

yentes si y sólo si A – M no serían verdaderas si N fuera falsa. El análisis es más débil porque decir que A – M no *serían* verdaderas si N fuera falsa no es decir que no *podrían* ser verdaderas si N fuera falsa, tal y como pide el análisis más fuerte. Es tan débil que no nos da un sentido genuino de «concluyente», lo cual no importa demasiado. Me parece que, en líneas generales, este análisis más débil es provechoso, y la teoría que apoyaré en el siguiente capítulo es muy parecida. Pero se diferencia de él en que no habla de razones, lo que es una ventaja porque parece posible que haya creencia justificada sin razones. Quizá sea posible que mi creencia de que tengo dolor esté justificada. Pero es difícil decir que yo la fundamento en razones, concluyentes o no. No la fundamento en ningún tipo de razones.

LA TEORÍA CAUSAL

A. I. Goldman propone un complemento causal a la definición tripartita [Goldman (1967)]. Un diagnóstico inicial de los contraejemplos de Gettier puede ser afirmar que es mera suerte el que la creencia justificada de Henry sea verdadera. Este diagnóstico no puede proporcionar por sí mismo una respuesta adecuada. No nos podemos limitar a estipular que no esté involucrada la suerte, porque todos nosotros la necesitamos en mayor o menor medida. Por ejemplo, el hecho de que un método fiable de formación de creencia nos proporcione ahora una creencia verdadera más bien que una falsa, como sucede a veces, estará en función de la suerte que tenga cada uno de nosotros. Por supuesto, el escéptico encuentra también un punto de apoyo en el hecho de que siempre esté implicada la suerte. Pero el diagnóstico puede sugerir una respuesta mucho mejor. La sugerencia de Goldman equivale a decir que lo que hacía que la creencia de Henry fuera verdadera en nuestro ejemplo no es lo que hacía que él la creyera. De modo que propone que, como cuarta condición para el conocimiento de que *p*, que la creencia de *a* en *p* sea causada por el hecho de que *p*. Con ello quedan excluidos los casos de Gettier porque en ellos es una pura coincidencia que la creencia sea verdadera. Queremos un vínculo entre creencia y verdad que evite que suceda tal cosa, y un vínculo causal parece atractivo a este respecto.

Por atractivo que sea este punto de vista, plantea varias dificultades. La primera es la dificultad en suponer que los hechos causen algo; son demasiado inertes como para afectar la manera en que funciona el mundo, incluso si el mundo es el mundo meramente mental

de las creencias. Después de todo, ¿qué son los hechos? Una primera idea es que los hechos son similares, si no idénticos, a las proposiciones verdaderas (lo que explicaría por qué no hay hechos falsos). Pero ¿pueden las proposiciones verdaderas causar algo? Ciertamente, los hechos (o las proposiciones verdaderas) parecen reflejar el mundo más que actuar sobre él. Los análisis más corrientes de la causación parecen estar en lo cierto al no permitir que sean causas más que los sucesos y, posiblemente, los agentes. En segundo lugar, hay un problema sobre el conocimiento del futuro; la sugerencia de Goldman parece implicar que nos encontramos en este punto con un caso de causación hacia atrás *(backward causation)* —el futuro causando el pasado— o que el conocimiento del futuro es imposible, ya que las causas no pueden ser posteriores a sus efectos. En tercer lugar, está el problema del conocimiento universal o, en líneas más generales, del conocimiento por inferencia. Mi creencia de que todos los hombres son mortales está causada, pero no por el hecho de que todos los hombres sean mortales; si hay hechos que causen tal creencia, son los hechos de que este hombre, ese hombre, etc., han muerto. Y lo que ha causado que esos hombres mueran no es el hecho de que todos los hombres mueran (lo que reivindicaría de nuevo al análisis causal, con una causa intermedia); más bien sucede que todos los hombres mueren porque esos hombres concretos (junto con otros muchos) lo hacen. ¿Cómo podría mostrar el análisis causal que yo sé que todos los hombres mueren?

Hay respuestas a algunas de estas críticas, por supuesto. Estamos más acostumbrados a hablar de hechos como causas de lo que nos permite la primera crítica. El hecho de que los filósofos no se hayan convencido de que comprenden la idea de que los hechos pueden ser causas no debe hacer que prohibamos cualquier apelación a la causación de hechos como filosóficamente injustificada. La segunda crítica también podría responderse si complicáramos la teoría y concediéramos que pudieran conocerse ciertos hechos cuando hecho y creencia son efectos distintos de una causa común. Sin embargo, la tercera crítica parece más intratable. La admisión de que los hechos pueden ser causas no hará más aceptable el supuesto de que los hechos universales pueden causar creencias universales. A pesar de todo ello, hay que decir que en la teoría causal hay algunos aspectos atractivos, y la teoría que voy a apoyar puede verse, de hecho, como una generalización a partir de ella.

2.4. OBSERVACIONES FINALES

Las diversas propuestas de la sección anterior se han presentado como si fueran añadidos al análisis tripartito, al haberse admitido que Gettier había mostrado la insuficiencia de tal análisis. Pero, entre ellas, podemos descubrir al menos una que podría considerarse como una defensa directa de ese análisis. Cualquier propuesta que equivalga a una nueva teoría de la justificación podría mostrar con éxito que, en el caso de Gettier, las creencias relevantes verdaderas no estaban justificadas en absoluto. La teoría causal podría decirnos que una creencia sólo está justificada cuando está causada (directa o indirectamente) por los hechos. En ese caso, adoptaría la ruta 1 (tal y como se distinguió en 2.2). (Algunas versiones de la propuesta de la fiabilidad también podrían verse bajo esta luz.) Realizando este movimiento, partiríamos de una teoría causal de la justificación; la teoría causal del conocimiento sería tan sólo una de sus consecuencias.

Un posible argumento contra la teoría causal de la justificación podría pretender que no tenemos garantía alguna de que sólo haya un modo en que nuestras creencias se justifican, y, en particular, ninguna razón real para suponer que cualquier modo aceptable debe ser causal y que todas las creencias justificadas de que *p* deben estar causadas por hechos relevantes. Ciertamente, aunque no quisiéramos admitir la existencia de hechos morales, no desearíamos, por ello, prohibir de antemano la posibilidad de que algunas creencias morales estuvieran justificadas. También podríamos dudar de la existencia de hechos matemáticos causalmente efectivos, sin desear decir que, por ello, ninguna creencia matemática puede estar justificada.

Sin embargo, es más importante advertir que este análisis causal de la justificación es falso porque niega la posibilidad de que una falsa creencia esté justificada. Una falsa creencia de que *p* no tiene ningún hecho de que *p* para causarla. Esta objeción sólo puede ser evitada si encontráramos un análisis de la justificación de las creencias falsas diferente del que se ofrece para las verdaderas. Y tal cosa no podría ser correcta. La justificación debe ser la misma para las creencias verdaderas y para las falsas, aunque sólo sea porque podemos preguntar y decidir si una creencia está justificada (por ejemplo, una creencia sobre el futuro) antes de decidir si es verdadera o falsa.

Esta crítica deja abierta la posibilidad de un tipo diferente de teoría causal, sobre las líneas que se han sugerido al final de 2.2. Con una teoría causal del conocimiento y la tesis de que una creencia está

justificada si y sólo si, en el caso de ser verdadera, produciría conocimiento, podemos proporcionar un análisis causal de la justificación que no sea vulnerable a la existencia de creencias falsas justificadas.

LECTURAS COMPLEMENTARIAS

Ensayos cruciales en este tipo de problemas son los de Gettier (1963), Dretske (1971), Goldman (1967) y Swain (1974).

Es posible que la primera discusión (y el rechazo) del análisis tripartito se encuentre en un diálogo de Platón: el *Teeteto* (201c-210d).

Shope (1983) analiza detalladamente la enorme producción generada en los últimos años por la percepción de los defectos del análisis tripartito. Por supuesto, hay muchos puntos de vista, con sus variantes respectivas, sobre el problema de Gettier que no he discutido todavía, incluyendo el del mismo Shope.

Muchos de los ensayos a los que se ha hecho referencia en el presente capitulo están recogidos en Pappas y Swain (1978), que contiene también una introducción analítica al área.

Prichard (1967) proporciona un análisis de las relaciones entre conocimiento, creencia, certeza y verdad, que difiere de una manera muy interesante del que se ha ofrecido.

Una cuestión importante que no se ha discutido es la de si el conocimiento implica la creencia. Para esto, véase Rang (1977).

Los ensayos de Gettier, Prichard y Woozley están recogidos en Phillips Griffiths (1967).

3. LA TEORIA CONDICIONAL DEL CONOCIMIENTO

3.1. LA TEORIA

Esta teoría, que debemos principalmente a Robert Nozick, parte de donde lo hacen otras: de los defectos del análisis tripartito desvelados por Gettiter. Nozick sugiere que la razón por la que consideramos que, en esos ejemplos, las creencias verdaderas justificadas no son conocimiento es que *a* las habría creído incluso si hubieran sido falsas. La razón por la que la creencia de Henry de que McEnroe es el campeón de este año era demasiado afortunada, o verdadera con demasiada fortuna, como para ser considerada conocimiento es que el camino por el que aquél llegó a esa verdad era tal que, incluso si su creencia hubiera sido falsa, hubiera acabado creyéndola. Nozick considera, por tanto, que, para que *a* crea que *p*, necesitamos que *a* no hubiera creído que *p* si *p* hubiera sido falsa.

Con ello tenemos las dos condiciones habituales:

1. *p*
2. *a* cree que *p*,

junto con

3. si *p* no fuera verdadera, *a* no creería que *p*.

Pero Nozick argumenta que, aunque este análisis puede enfrentarse a los ejemplos que nos ofrece Gettier, existen otros ejemplos semejantes que escaparían a la formulación presente. Hay dos maneras en las que puede ser una coincidencia que la creencia de *a* sea verdadera, y necesitamos eliminar las dos. La primera es aquella en la que *a* creería *p*, incluso cuando *p* fuera falsa. La adición de la cláusula 3 intenta neutralizar este hecho. La segunda es que podrían existir circunstancias ligeramente distintas en las que *p* continuara siendo verdad, sin que *a* creyera en ella. Hay muchos ejemplos de coincidencia feliz en ambos sentidos. Supongamos que creo que hay un coche de policía ahí

afuera porque puedo oír una sirena. De hecho hay un coche de policía, pero la sirena que oigo proviene del aparato de música de mi hijo en la habitación de al lado. No sé que haya un coche de policía en la calle, por dos razones diferentes. En primer término, hubiera tenido la creencia aunque el coche no hubiera estado en ese lugar. En segundo término, no hubiera tenido la creencia si mi hijo no hubiera estado escuchando música, por más que el coche hubiera continuado estando a la puerta de la calle. Eliminamos esta segunda forma en la que una creencia verdadera podría ser demasiado afortunada como para contar como conocimiento, añadiendo una cuarta cláusula a las otras tres:

4. si, en circunstancias diferentes, *p* todavía fuera verdadera, *a* todavía creería que *p*.

Estas cuatro condiciones, que constituyen la teoría condicional del conocimiento, pueden simbolizarse de la manera siguiente:

1. p
2. Cap
3. $\sim p \rightarrow \Box \sim Cap$
4. $p \Box\!\!\rightarrow Cap$

donde el cuadrado con flecha «$\Box\!\!\rightarrow$» se usa para simbolizar la construcción subjuntiva en castellano «Si fuera el caso que..., sería el caso que...». La teoría así constituida es un intento de articular la intuición de que, para que una creencia sea conocimiento, debe ser particularmente sensible a la verdad de la proposición creída. Por expresarlo en términos del propio Nozick, debe *seguir el rastro* a la verdad, en el sentido de que, si la proposición fuera verdadera en circunstancias diferentes, todavía sería creída y, si no fuera verdadera, no lo sería.

3.2. ALGUNOS COMENTARIOS

a) El requisito de que la creencia de que *p* debe seguir el rastro a la verdad de *p* es el requisito de que las dos primeras cláusulas de la teoría deben estar relacionadas de alguna manera. Es algo similar a la manera en que funcionaba la teoría causal, aunque en tal caso la relación que se requería era, específicamente, una relación causal. La teoría condicional es menos exigente, de modo que escapa a algunas de las dificultades a las que se enfrentaba la teoría causal. Pero incluye la

teoría causal como un caso especial, dado que podríamos pensar que, si el hecho de que *p* causa la creencia de *a* en *p*, en ese caso los dos condicionales subjuntivos serán verdaderos (pero no viceversa). De modo que la teoría condicional es una generalización de la teoría causal. Mientras que la teoría causal sugiere que sólo hay una manera en la que las creencias verdaderas justificadas pueden llevar a ser conocimiento (el que sean causadas por los hechos), la teoría condicional está dispuesta a considerar cualquier manera, causal o no, que preserve la verdad de los dos condicionales subjuntivos.

De hecho, la teoría condicional adopta algunas de las mejores intuiciones de algunas de las teorías que, en el capítulo anterior, considerábamos que eran defectuosas. Por ejemplo, está muy cerca de la elaboración que hizo Drestske del punto de vista de las «razones concluyentes».

b) Considerábamos si la teoría causal del conocimiento descansaba sobre (o nos permitía acceder a) una teoría causal de la justificación. Es obvio que podríamos preguntar lo mismo respecto a la teoría condicional. ¿Podemos ofrecer una definición condicional de CJ*ap* como

$$\mathrm{CJ}ap \equiv (p \mathrel{\Box\!\!\rightarrow} \mathrm{C}ap \ \& \ {\sim}p \mathrel{\Box\!\!\rightarrow} {\sim}\mathrm{C}ap)?$$

Esto equivaldría a mantener que una creencia justificada es aquella que sigue el rastro a la verdad. Pero surgen dificultades del mismo tipo. Una creencia falsa puede estar justificada. Entonces, CJ*ap* es consistente con (C*ap* & ~*p*). Pero (C*ap* & ~*p*) es inconsistente con ($p \mathrel{\Box\!\!\rightarrow} \mathrm{C}ap \ \& \ {\sim}p \mathrel{\Box\!\!\rightarrow} {\sim}\mathrm{C}ap$); una creencia falsa no sigue el rastro de la verdad. De modo que la definición condicional de la justificación falla.

A pesar de todo, una teoría semejante debería ser atractiva. Con toda seguridad, alguien que considere que su creencia de que *p* está justificada está muy cerca de aceptar que su creencia sigue el rastro a la verdad de *p*. Podemos, si lo deseamos, restaurar la posibilidad de una teoría condicional de la justificación por medio de un movimiento similar al que hicimos en 2.4. La sugerencia en ese punto era la de derivar el análisis de la justificación a partir del análisis del conocimiento: *a* está justificado por creer que *p* si y sólo si en ciertas circunstancias *a* sabría que *p*. A modo de ilustración, podemos considerar el siguiente ejemplo:

$$CJap \equiv (p \;\Box\!\!\rightarrow Sap)$$

La expresión crucial «en ciertas circunstancias» debe ser leída de la manera más simple posible, como «si *p* fuera verdad». Si pudiéramos disponer de una teoría de este tipo, podríamos aceptar de una manera indirecta la posibilidad de una teoría condicional de la justificación.

c) Otro aspecto que, a mi parecer, está en favor de la teoría es que comienza a dar cierto sentido teórico al sentimiento intuitivo de que lo incorrecto en los ejemplos de Gettier era la suerte excesiva que en ellos se incorporaba. La teoría nos da el siguiente análisis de cuándo una creencia es verdadera por mera suerte: la creencia de *a* es verdadera por suerte en la medida en que *a* la hubiera continuado creyendo aunque fuera falsa, o aunque hubiera sido verdadera bajo ciertas condiciones diferentes. La importancia de este análisis resultará obvia en la sección siguiente.

d) La teoría parece poseer algunos recursos con los que explicar los vínculos entre certeza y conocimiento. Alguien que pretende que sabe que *p*, pretende que no creería en *p* si *p* no fuese verdadera, y que creería en *p* cuando *p* fuera verdadera. No podría tener tal pretensión si no estuviera seguro de *p*. Lo que había perdido el mal estudiante del ejemplo que mencionábamos anteriormente era la confianza en que sus creencias estuvieran rastreando la verdad; consideraba que, aunque creía que *p*, era por lo menos tan probable que estuviera en lo cierto como que no lo estuviera. Por tanto, la teoría analiza la certeza necesaria para las declaraciones de conocimiento como la creencia en que se satisfacen los dos condicionales subjuntivos. Utiliza este análisis para explicar el hecho, que en caso contrario resultaría difícilmente explicable, de que el mal estudiante sepa sin que pueda declarar que sabe.

3.3. EL PRINCIPIO DE CIERRE Y EL PRIMER ARGUMENTO ESCÉPTICO

La teoría condicional del conocimiento puede mostrar que no sabemos que no somos cerebros en una cubeta, dado que falla una condición necesaria para tal conocimiento, la tercera: que, si no fuera verdad que no somos cerebros en una cubeta, no creeríamos que no lo

somos. La condición no se cumple porque, si fuéramos cerebros en una cubeta (a los que se les provocan artificialmente experiencias), todavía creeríamos que no lo somos. Por tanto, si la teoría condicional del conocimiento está en lo cierto, no sabemos que no somos cerebros en cubetas. Una condición necesaria para tal conocimiento no se cumple.

A pesar de ello, el análisis de Nozick puede usarse para mostrar que el lector sabe que está sentado leyendo un libro de filosofía (por favor, siéntese usted antes de continuar). En este caso se satisfacen las cuatro condiciones: es cierto que el lector está sentado, etc.; el lector cree que está sentado, etc.; si no estuviera sentado, etc., no creería que lo está; si estuviera sentado, etc., creería que lo está.

El análisis muestra también, de una manera semejante, que sabemos que, si estamos sentados leyendo un libro de filosofía, no somos cerebros en una cubeta. De donde se sigue que, de acuerdo con la teoría condicional del conocimiento, es posible saber que p y saber que p implica q sin saber que q. Esta conclusión ¿no rompe directamente el principio de cierre $[(Sap \ \& \ Sa\ (p \rightarrow q)] \rightarrow Sag$? Lo hace; pero Nozick es capaz de mostrar de una manera más general que, según este análisis, el principio no es aceptable. La explicación descansa en los aspectos más generales de una teoría sobre las condiciones en las que condicionales subjuntivos como 3 y 4 son verdaderos.

La teoría usa la noción de «mundo posible» para proporcionar un análisis de las condiciones de verdad de los condicionales subjuntivos. Un mundo posible debe ser considerado como una forma completa en la que el mundo podría haber estado, un modo en el que el mundo podría haber existido. La idea de que el mundo podría haber sido diferente en ciertos respectos se considera la idea de que hay un mundo posible que difiere del actual en ese respecto (y, posiblemente, otros también). Los mundos posibles varían en su grado de parecido con el mundo real. Algunos están muy cerca del nuestro, otros son mucho más lejanos. Pero, por dos razones, puede que no sea posible ordenar mundos de acuerdo con su cercanía al mundo real. En primer lugar, las nociones de cercanía, parecido y semejanza son demasiado imprecisas como para apoyar el tipo de juicios comparativos precisos que exigiría semejante ordenamiento. En segundo lugar, siempre puede ser razonable esperar que, para cada mundo posible, podamos encontrar otro que se parezca en el mismo grado al real. Por estas dos razones, es preciso que pensemos en términos no de mundos posibles individuales, sino de grupos de mundos posibles, con todos los mundos de un mismo grupo estando a la misma distancia del mundo real.

Si hay una ordenación, aunque sea vaga e imprecisa, debe serlo de grupos de mundos igualmente similares más bien que de mundos individuales.

Nozick da un análisis inicial de las condiciones de verdad para el condicional subjuntivo $p \Box\!\!\rightarrow q$:

> $p \Box\!\!\rightarrow q$ es verdadero en el mundo real si y sólo si $p \rightarrow q$ es verdadero en todo un rango de grupos de mundos razonablemente cercanos al mundo real.

(Ésta no es la teoría final, sino una primera aproximación.)

Aunque esta teoría se expresa formalmente, podemos darle fácilmente un apoyo intuitivo. Supongamos que queremos saber las circunstancias en las que estaría justificada la aseveración del condicional «Si la Sra. Thatcher hubiera esperado más, habría perdido las elecciones». Lo que hacemos es representarnos la situación actual, alterándola hasta el punto en que la Sra. Thatcher retrasa las elecciones. Por supuesto, habría que alterar otras muchas cosas, dado que la fecha de unas elecciones no es un hecho aislado (no hay hechos aislados, y ésa es la razón por la que no puede haber un mundo similar al nuestro en todos los respectos excepto en un hecho). Tendríamos que cambiar también todas las cosas que no hubieran sucedido si las elecciones se hubieran retrasado, por ejemplo, la edad de la Sra. Thatcher en la fecha de las elecciones. Luego, manteniendo constantes las demás cosas, en la medida en que sea posible, formamos un juicio sobre lo que hubiera tenido una mayor probabilidad de suceder. Si consideramos que lo más probable hubiera sido una derrota del partido conservador, aseveramos el condicional subjuntivo. En caso contrario, no.

La teoría formal completa esta aproximación informal diciendo que consideramos los grupos de mundos más cercanos en los que el antecedente p es verdadero y nos preguntamos si, en esos mundos, q es también verdadero. Si q es probable dado que p, entonces p & q es más probable que p & $\sim q$, y, por tanto, en los mundos más cercanos en los que p es verdadero, q será también verdadero. Por supuesto, habrá mundos más lejanos en los que tendremos p & $\sim q$, pero no importa, puesto que estamos diciendo lo que sucedería (más probablemente) no lo que podría (dentro de los límites de la mera posibilidad) suceder.

Esta distinción entre lo que sucedería y lo que podría suceder es crucial para lo que sigue. Por considerar el ejemplo de David Lewis

[D. Lewis (1973)], «si los canguros no tuvieran cola, se caerían»; es verdad, por supuesto, que siempre *es posible* que no cayeran: es posible que un gobierno con ganas de fomentar el turismo les diera muletas. Pero nada de eso afecta al hecho de que caerían, dado que, en los mundos más similares a los nuestros, los canguros sin cola caen y nadie les regala muletas.

El análisis de Nozick puede usarse ahora de dos maneras. En primer lugar, puede reforzar nuestras intuiciones de que el lector está sentado, leyendo, etc., que sabe que si está sentado y leyendo, etc., no es un cerebro en una cubeta, y que no sabe que no es un cerebro en una cubeta. Con la teoría condicional del conocimiento, y la explicación precedente de las cláusulas 3 y 4, puede decirse que el lector sabe que está sentado leyendo porque (3) en los mundos más cercanos en los que no está sentado leyendo no cree que lo está, y (4) en los mundos más cercanos en los que está sentado leyendo cree que lo está. A pesar de ello, el lector no sabe que no es un cerebro en una cubeta porque no es el caso que (3) en los mundos más cercanos en los que es un cerebro en una cubeta él crea que es un cerebro en una cubeta. Pongo de nuevo de relieve que sólo se tienen en cuenta los mundos más *cercanos* en los que el antecedente es verdadero. Se admite que sea posible, por ejemplo, que el lector crea que está sentado leyendo cuando en realidad no lo está. Pero un mundo semejante está mucho más lejos del actual que un mundo en el que, cuando está sentado y leyendo, cree que lo está. Estamos interesados no en lo que *podría* ser el caso (por supuesto, todo es posible), sino en lo que *sería* el caso, y la teoría señala que ése es el centro de interés al tomar en consideración sólo a los mundos relevantes que están más cerca del nuestro.

La teoría proporciona también una contraprueba directa del principio de cierre. Para ello tiene dos caminos: o bien dar la descripción de un mundo para el que el principio falla y probar formalmente que la descripción es consistente, o bien utilizar un método más informal y apelar a un contraejemplo. Ya tenemos uno de esos contraejemplos: el caso en el que a = el lector, p = el lector está sentado leyendo, q = el lector no es un cerebro en una cubeta. Pero el ejemplo es discutible y necesita del apoyo de una explicación, en términos de la teoría condicional, de cómo son posibles tales contraejemplos. Eso es lo que ahora ya podemos hacer.

Una tarea previa e importante es distinguir el principio de cierre, que es falso, del caso siguiente de *modus ponens*, al que debemos conceder mucha más confianza:

$$[Sap \,\&\, (Sap \rightarrow Saq)] \rightarrow Saq.$$

En la medida en que mantengamos separados ambos principios, no es difícil darse cuenta de que el principio de cierre no funciona. La razón estriba en los condicionales subjuntivos y los mundos posibles que se consideran relevantes para la evaluación de esos condicionales como verdaderos o como falsos. Los mundos relevantes para evaluar la parte izquierda del condicional del principio de cierre son el grupo de mundos más cercano en el que p es verdadera, el grupo más cercano en el que p es falsa, el grupo más cercano en el que $p \rightarrow q$ es verdadera (que puede ser diferente) y el grupo más cercano en el que $p \rightarrow q$ es falsa. Los mundos relevantes para evaluar la parte derecha del principio de cierre son el grupo más cercano en el que q es verdadera y el grupo más cercano en el que q es falsa, es decir, grupos de mundos completamente diferentes, uno de los cuales puede estar, como en el ejemplo que considerábamos, mucho más alejado de nuestro mundo real que cualquiera de los primeros cuatro grupos. Por ello, no es probable que un conjunto de observaciones sobre los primeros cuatro grupos pudiera establecer alguna restricción sobre la naturaleza de los dos últimos grupos, por lo que queda un espacio enorme para un contraejemplo que invalide el principio de cierre. Cuanto más distante sea uno de esos últimos grupos, más probable será que podamos construir un contraejemplo; como quedaba claro en el propio ejemplo que considerábamos anteriormente, en el que $\sim q$ = el lector es un cerebro en una cubeta.

A modo de ilustración final, podemos aplicar la refutación de Nozick al argumento escéptico de Descartes sobre el sueño, y, así, es posible que nos vaya mejor que a Descartes en la última página de las *Meditaciones*. Sean a = el lector, p = el lector está sentado leyendo, q = el lector no está en el lecho soñando que está sentado leyendo. ¿Es posible que sean verdad Sap, $\sim Saq$ y $Sa(p \rightarrow q)$? Sí. Saq es falsa porque, en el grupo más cercano de mundos en el que q es falsa (el lector está en la cama leyendo, etc.), el lector todavía cree que es verdadera. Sap es verdadera porque en el grupo de mundos más cercanos en los que el lector está sentado y leyendo cree que está sentado y leyendo, y los mundos más cercanos en los que no está sentado y leyendo son o bien mundos en los que está en pie/arrodillado/acostado leyendo, o bien mundos en los que está sentado cosiendo/mirando la televisión, etc., pero no mundos en los que suceda que el lector está dormido y soñando que está sentado leyendo. De modo que el lector Puede saber que está sentado leyendo, aunque no sepa que no está dorado en la cama, soñando que está sentado leyendo.

3.4. ¿HA REFUTADO NOZICK AL ESCÉPTICO?

Nuestro objetivo en la sección precedente era mostrar que la preocupación por las virtudes y los defectos de la condición tripartita puede ser de genuino provecho filosófico. La teoría condicional es atractiva de propio derecho como un análisis verosímil del conocimiento que escapa a las objeciones del tipo de Gettier. Tiene, también, la virtud añadida de que bloquea un tipo habitual de estrategia escéptica, y lo hace mostrando, además, la razón por la que esa estrategia parece tan atractiva.

De acuerdo con Nozick, sin embargo, todos los argumentos escépticos descansan en el principio de cierre y la invalidación de tal principio sirve, por tanto, de respuesta general al escepticismo [Nozick (1981), p. 204]. Aparte de que esta sugerencia sea implausible en líneas generales, hay un argumento escéptico que no puede rechazar de esa manera. Se trata del argumento del error (1.2).

El argumento es que nosotros, o los demás, hemos cometido errores en el pasado o los cometeríamos en circunstancias en las que, en la medida en que podemos distinguirlas, no son relevantemente distintas de las circunstancias presentes. Por tanto, no podemos, admitiendo que nosotros o ellos no sabíamos en un momento determinado anterior, insistir en que sí sabemos ahora, dado que eso sería hacer declaraciones distintas en circunstancias que no muestran ninguna diferencia relevante.

Este argumento nos obligaba a admitir, en primer lugar, que no sabemos que no somos cerebros en cubetas. Nozick necesita de esta conclusión para poder generar el problema escéptico que trata de disolver, dado que su escéptico argumenta que el consecuente Saq de PC^S es falso en este caso y, por ello, admitiendo la verdad de $Sa(p \rightarrow q)$, que Sap es falsa. Nozick desea considerar este argumento escéptico, a la vez que admite la legitimidad de su primer movimiento, y es el argumento a partir del error el que fuerza tal admisión. De modo que, sólo si admite la fuerza del argumento del error puede pensar que llega a algún sitio al refutar PC^S.

Nozick podría decir que su teoría condicional prueba también el mismo extremo: que no sabemos que no somos cerebros en una cubeta. Y está en lo cierto; en ese caso, nuestra creencia no seguiría el rastro de la verdad. Pero si no tuviéramos razones independientes para aceptar esa conclusión, como las que proporciona el argumento del error, lo consideraríamos como algo *en contra de* su teoría, al mostrar que no conocemos las cosas más centrales y obvias, como que no

somos cerebros en cubetas, que hay un mundo material o que el universo existe desde hace más de cinco minutos. De modo que no puede arrojar por la borda el argumento del error y confiar completamente en la teoría condicional para mostrar que S*aq* es, en este caso, falsa.

En segundo lugar, hay obviamente algo de correcto en un argumento que mantiene, como hace el argumento del error, que, por el mismo tipo de razones por el que el lector no sabe que no es un cerebro en una cubeta, tampoco sabe, por ejemplo, que *The Times* se publicará mañana ni que está sentado leyendo. Ciertamente, es inverosímil suponer, ateniéndonos al ejemplo de Russell [Russell (1921), p. 159] que yo sepa lo que hice ayer, sin saber que el mundo no comenzó a existir hace cinco minutos repleto de residuos arqueológicos y otros muchos restos del pasado. Nozick desea mostrarnos que podemos suponer tal cosa; pero no estoy seguro de que mi deseo sea aceptar tal supuesto (o de que deba desear aceptarlo). El argumento del error tiene, en ese punto, una consistencia que le hace plausible, mientras que lo que Nozick considera el punto fuerte de su propio argumento comienza a parecer su punto débil. (Esta observación podría reivindicarse como la base de un argumento general contra la teoría condicional del conocimiento.)

En tercer lugar, el argumento del error no puede ser atacado por descansar en PC^S. La razón de esto es independiente de la validez de PC^S. Incluso aunque fuera válido, no hay razón para suponer que el argumento descansa sobre tal principio. De cualquier modo, Nozick no puede mantener que sí lo hace, porque su posición descansa en la prueba independiente de que el lector no sabe que no es un cerebro en una cubeta. Sin semejante prueba, mantendríamos el principio en contra de la teoría condicional del conocimiento, que ni siquiera es capaz de mostrar que sabemos que no somos cerebros en una cubeta. Con dicha prueba, Nozick puede encontrar y encuentra un apoyo independiente para su propia teoría; desea decir que la teoría está en lo cierto en ese extremo particular. Pero Nozick pierde ese apoyo independiente y, con él, una gran dosis de verosimilitud si el argumento del error descansa en PC^S que, de acuerdo con su propia teoría, no es un principio válido.

En cuarto lugar, el argumento no puede rechazarse sobre el razonamiento similar de que parte del hecho de que podríamos equivocarlos, tratando de alcanzar directamente la conclusión de que no sábelos. Tal tipo de argumentación, como Nozick puede mostrar, sería JOválido. Pero el argumento no sigue este desarrollo directo y falaz, aÒK) una ruta indirecta, utilizando el principio de universalizabilidad.

3.5. INTERNALISMO Y EXTERNALISMO

Hasta ahora lo que tenemos es que el análisis de Nozick tiene éxito contra un argumento escéptico, pero no contra otro. ¿Es una conclusión que Nozick deba aceptar sin reticencias? ¿Debe abandonar cualquier pretensión de éxito total y limitarse a señalar el éxito incontrovertible pero parcial?

Todavía hay una queja que Nozick podría hacer y que probablemente hiciera. Podría decir que su concepción del conocimiento es una concepción *externalista*, mientras que nuestro argumento escéptico a partir del error es un argumento *internalista*. Si el externalismo es una actitud sensata, el argumento del error será irrelevante, dado que no hace sino elaborar un punto de vista defectuoso (aunque tradicional) sobre la epistemología. Si hace algo más, sólo muestra cómo ese punto de vista defectuoso conducirá al escepticismo.

¿Qué se quiere decir al hablar de concepciones «internalistas» y «externalistas»? La respuesta a esta cuestión puede darla un ejemplo. La teoría causal del conocimiento, que define *«a* sabe que *p»* como equivalente a

1. *p* es verdadera
2. *a* cree que *p*
3. la creencia 2 de a está justificada
4. la creencia 2 de *a* está causada por el hecho de que *p* es verdadera

es una concepción externalista, porque *a* podría ser completamente incapaz de reconocer o de señalar a la condición 4, si se le preguntara si sabe que *p*. En este caso, el externalismo dice que, en la medida en que la condición 4 se mantenga de hecho y se supongan dadas las condiciones 1–3, a sabe que *p*, independientemente de si es o no capaz de señalar o comprender tal condición. El internalista pretenderá que para que la cláusula causal pueda convertir la creencia justificada en conocimiento, es preciso no sólo que sea verdad, sino que *a* crea que lo es. Así pues, el internalista añadiría

5. *a* cree 4.

Hay argumentos a favor del externalismo y argumentos a favor del internalismo. El externalista puede señalar lo difícil que va a ser para el internalista el proporcionar un análisis satisfactorio del conocimien-

to. Ciertamente, podría decir, si añadimos la cláusula 5, debemos añadir también

6. la creencia 5 de a está justificada,

y, después, es de presumir que necesitara añadir nuevas cláusulas del tipo de

7. la creencia 5 de a está causada por el hecho de que 5 es verdadera.

Hemos generado un regreso al infinito, lo que quiere decir que el internalismo está condenado al escepticismo. Lo que es más, el regreso no depende de los elementos causales del ejemplo usado. Podríamos crear el mismo regreso si comenzáramos por la concepción tripartita tradicional y añadiéramos, basándonos en consideraciones interaalistas,

4. *a* cree 3

y, después, presumiblemente requeriríamos

5. la creencia 4 de a está justificada

y, después,

6. *a* cree 5,

y así sucesivamente.

El internalista puede responder señalando lo enorme que es nuestra intuición natural en favor de la concepción internalista. Supongamos que elaboramos el ejemplo causal; para el conocimiento, se requiere que la cuarta cláusula sea verdadera, pero no que *a* tenga indicio alguno de que lo es. ¿No muestra eso que, por todo lo que *a* sabe, no sabe que *p*? ¿Y cómo puede saber que *p* cuando, por todo lo que sabe, no lo sabe?

En mi opinión, ninguno de esos argumentos es efectivo para destrozar a su opositor. El primero se limita a señalar las dificultades con el escepticismo; el internalista aceptaría eso y diría que tales dificultades han de afrontarse. El segundo parece equivaler más a un enunciado de la posición externalista que a un argumento independiente con-

tra el internalismo. De hecho, dudo de que pueda existir un argumento concluyente en favor de cualquiera de estos puntos de vista; éstos son tan distintos que cualquier argumento dará por sentada la cuestión principal.

La posición de Nozick pretende ser externalista [Nozick (1981), pp. 265-268 y 280-283]. Las condiciones 3 y 4 son evidencia obvia de externalismo, y no hay ninguna sugerencia de que sea necesario añadir nada como

5. $Ca(\sim p \rightarrow \sim Cap)$.

En tanto que externalista, ¿no estaría justificado para dejar a un lado el argumento del error como expresión irrelevante de internalismo? No lo creo, por dos razones. En primer lugar, de nuevo, Nozick necesita el argumento, internalista o no, como un apoyo independiente de algo que, si careciera de él, parecería una consecuencia contraintuitiva de la teoría. En segundo lugar, al considerar que el argumento es necesario, Nozick acepta tácitamente que su teoría no es tan puramente externalista como podría parecer en un principio, por lo que no puede pretender inmunidad para un ataque con la única excusa de que es un ataque internalista.

Volveremos al contraste entre internalismo y externalismo en 9.3, y lo examinaremos con mayor detalle. Mientras tanto, ¿dónde nos ha dejado nuestra discusión de la teoría condicional? Todavía no tenemos ninguna respuesta al argumento escéptico del error. Hemos mencionado, sin adoptarlas, dos posiciones que parecen prometedoras en este respecto: externalismo y antirrealismo. Sin embargo, aquellos que no encuentran ningún atractivo en estas posiciones deberán buscar más si quieren decir que existen cosas tales como el conocimiento o la creencia justificada. El argumento del error volverá a aparecer, como una amenaza persistente. Aventuraré mi propia respuesta en el capítulo final.

LECTURAS COMPLEMENTARIAS

La teoría condicional y sus consecuencias para el escepticismo se exponen en Nozick (1981), especialmente en pp. 172-178 y 197-227.

Stroud (1984, cap. 2) considera un argumento similar contra el intento del escéptico de generalizar a partir del hecho de que no sabemos que no somos cerebros en una cubeta. Prosigue la discusión en el capítulo 7 de su libro.

En Garret (1983) se ofrecen contraejemplos a la teoría condicional; Gordon (1984) intenta una respuesta. Shope (1984) ofrece más contraejemplos.

D. Lewis (1973) ofrece un análisis pionero de las condiciones de verdad para los condicionales subjuntivos (que él denomina «contrafácticos») en términos de mundos posibles. Discuto una diferencia importante entre D. Lewis y Nozick en Dancy (1984b).

Al final del capítulo 9 se dan referencias para el debate entre internalismo y externalismo.

PARTE II

JUSTIFICACIÓN

4. FUNDAMENTALISMO

4.1. FUNDAMENTALISMO CLÁSICO

Quizá la posición más influyente en epistemología sea la que denominaré fundamentalismo clásico. La discusión de la justificación, de lo que sea una creencia justificada, comienza con esta teoría. Otras teorías se describirán en términos de su relación o su diferencia con ella. Nos ofrece una imagen persuasiva de los propósitos de la epistemología. En pocas palabras, equivale a una definición de la tarea epistemológica.

El fundamentalismo clásico divide nuestras creencias en dos grupos: las que necesitan el apoyo de otras y las que pueden apoyar a otras sin necesitar ellas mismas ningún tipo de fundamentación. Estas últimas constituyen nuestros fundamentos epistemológicos; las primeras, la superestructura construida sobre esos fundamentos.

La distinción entre fundamentos y superestructura, entre creencias básicas y creencias no básicas, es una distinción estructural. Pero el fundamentalismo clásico da contenido a la distinción añadiendo que las creencias básicas son creencias relativas a la naturaleza de nuestros propios estados sensoriales, o nuestra experiencia inmediata. Tales creencias descansan en sus propios pies, sin apoyarse en otras. Otras creencias necesitan de apoyos, y los obtienen de las creencias sobre nuestros propios estados sensoriales.

El fundamentalismo clásico, de este modo, da expresión a una asunción central del empirismo, el punto de vista de que todo nuestro conocimiento deriva de la experiencia. Lo hace así, insistiendo en que una creencia que no trata de nuestros propios estados sensoriales (experiencia inmediata) sólo puede justificarse apelando a creencias que tratan de ellos.

¿Cómo es posible que las creencias sobre nuestros estados sensoriales presentes no necesiten de ningún apoyo en otras, mientras que todas las demás creencias precisan de tal apoyo? La respuesta hay que buscarla en el tercer elemento del fundamentalismo clásico: el supuesto de que las creencias sobre nuestros estados sensoriales presentes son

infalibles. Es por ello por lo que pueden desempeñar el papel que se les adjudica en esta forma de empirismo; las creencias sobre nuestros estados sensoriales presentes pueden ser nuestra base —pueden mantenerse sobre sí mismas y sustentar a las demás— porque son infalibles.

Ya podemos ver qué es la epistemología, de acuerdo con el fundamentalismo. Es un programa de investigación que trata de mostrar cómo es posible que nuestras creencias sobre el mundo externo, sobre la ciencia, sobre el pasado y el futuro, sobre las otras mentes, etc., puedan justificarse sobre una base que está restringida a las creencias infalibles sobre nuestros propios estados sensoriales. Se sugiere que, si podemos hacer tal cosa, las exigencias de la epistemología se verán satisfechas. En caso contrario, nos aguarda el escepticismo.

En este capítulo y en el siguiente examinaremos el fundamentalismo clásico en algún detalle, y encontraremos razones para rechazarlo prácticamente en su totalidad. Pero antes debemos examinar los motivos y los argumentos que llevan o han llevado a los filósofos a seguir esa dirección. Ya hemos visto que el fundamentalismo clásico es una expresión de empirismo. Pero, como veremos, hay otras expresiones de empirismo. ¿Por qué debemos escoger ésa particularmente?

PROBABILIDAD Y CERTEZA

C. I. Lewis, el fundamentalista clásico más eminente del presente siglo, mantuvo que «a menos que algo sea cierto, nada más puede ser ni siquiera probable» [véase C. I. Lewis (1952)]. Este punto de vista puede ser entendido mejor si nos aproximamos a él con (un ligero) conocimiento del cálculo de probabilidad. En este cálculo, la probabilidad se evalúa siempre con relación a la evidencia. No preguntamos cuál es la probabilidad absoluta de una hipótesis h [escrito como $P(h)$]. En lugar de ello, nos preguntamos sobre la probabilidad condicional de h dada la evidencia e [escrito como $P(h/e)$]. La probabilidad de h dada e se expresa por medio de correlaciones que normalmente están en la escala del 0 al 1. Si $P(h/e) = 0$, dada e, es completamente seguro que h es falsa. Si $P(h/e) = 1$, dada e, es seguro que h es verdadera. Si, dada e, $P(h/e) = 0.5$, es tan probable que h sea verdadera como que sea falsa, puesto que, en el cálculo, $P(h/e) + P(\sim h/e) = 1$.

Lo más importante es que, al evaluar la probabilidad de h dada e, no cuestionamos e; temporalmente asumimos que e es cierta y pasamos por alto la posibilidad de que e no fuera verdadera. Pero la misma e tiene una probabilidad con relación a alguna otra evidencia e', y así

indefinidamente. Y, a menos que encontremos al final una proposición o un conjunto de evidencias *e"* que, de algún modo, tenga por si misma la probabilidad 1, toda esa serie de probabilidades no tendrá nada en qué apoyarse. Necesitamos encontrar algo cierto que pueda funcionar como la evidencia no cuestionada, por apelación a la cual evaluar las probabilidades de las demás cosas.

En este argumento se *sugiere* que una proposición con probabilidad 1 es *cierta*. Pero certeza e infalibilidad son cosas distintas, y estamos tratando de explicar una teoría que considera infalibles a sus propias creencias básicas. El paso de una a otra no es difícil. Si una proposición que es cierta tiene la probabilidad 1, no hay ninguna posibilidad de que una creencia en esa proposición sea falsa; de modo que tal creencia será infalible.

Hay algo de extraño en este argumento, que comienza insistiendo en que hablamos sólo de probabilidad relativa a la evidencia, y acaba por hablar de que una proposición tiene una probabilidad 1 de propio derecho. Los teóricos de la probabilidad eluden esta dificultad definiendo la probabilidad absoluta en términos de la probabilidad relativa: dicen que la probabilidad absoluta de h = la probabilidad de h relativa a una tautología. [$P(h) = P(h/q \text{ v } \sim q)$.] Es dudoso que esta maniobra sea algo más que un mero recurso técnico.

EL ARGUMENTO DEL REGRESO DE JUSTIFICACIONES

Todos están de acuerdo en que algunas de nuestras creencias están justificadas por su relación con otras creencias. Normalmente, se cree que esa relación es inferencial; una creencia se infiere de otra u otras. De modo que mi creencia en que, rascando la cerilla, ésta se encenderá está justificada inferencialmente. La he inferido (por supuesto, no de una manera consciente) a partir de otras creencias, probablemente creencias sobre ocasiones similares en el pasado.

El argumento del regreso de justificaciones supone que, tanto como las creencias justificadas inferencialmente, deben existir algunas creencias que se justifiquen no inferencialmente. Su intuición fundamental queda clara si suponemos que la inferencia es básicamente asunto de pasar de premisas a conclusión siguiendo una ruta aceptable. Si las premisas carecen de justificación, no habrá justificación alguna para la conclusión —al menos, no por *esa* inferencia—. Podemos suponer, entonces, que sólo las creencias justificadas pueden justificar las otras. Ésta es la idea que genera la regresión.

Supongamos que toda justificación es inferencial. Cuando justificamos la creencia A, apelando las creencias B y C, todavía no hemos mostrado que A esté justificada. Sólo hemos mostrado que A está justificada si lo están B y C. La justificación por inferencia es sólo justificación condicional; la justificación de A está condicionada a la justificación de B y C. Pero, si toda justificación inferencial es condicional en este sentido, no hay nada de lo que pueda decirse que está justificada realmente, de un modo no condicional. Para cada creencia que intentamos justificar, siempre habrá una creencia ulterior de cuya justificación dependerá la justificación de la primera, y, dado que éste es un regreso al infinito, no habría ninguna creencia que estuviera justificada más que de un modo condicional.

La única escapatoria a esta conclusión es suponer que la cadena de justificaciones, en vez de extenderse hasta el infinito, se retuerce y se une a sí misma en cierto punto, formando una especie de círculo. Pero con ello no se arreglan las cosas, porque todavía será el caso que la justificación de todos los miembros de este anillo es condicional: nunca podrá tener éxito en la eliminación de la condicionalidad.

El argumento del regreso de las justificaciones, por tanto, nos lleva a suponer que debe haber alguna justificación que no sea inferencial siempre que consideremos inadmisible la consecuencia escéptica de que ninguna creencia está justificada realmente. Y la pretensión de que hay dos formas de justificación, inferencial y no–inferencial, es el núcleo de cualquier forma de fundamentalismo en la teoría de la justificación.

Aparte de la rendición directa, hay una diversidad de respuestas posibles a este argumento. En 9.1 se mencionará una muy importante. Mientras tanto, debemos preguntar si esta regresión es tan peligrosa como podría parecer. No todo regreso al infinito es vicioso. Algunos son *virtuosos*, es decir, podemos convivir con ellos y no es preciso encontrar una manera de detenerlos. Por ejemplo, la regresión generada por la observación de que siempre hay un punto entre cualquier par de puntos puede ser virtuosa, incluso aunque la consideremos con relación a puntos del tiempo más bien que del espacio. Del mismo modo, podríamos aceptar la regresión temporal causada por el supuesto de que para cada momento en el tiempo siempre hay un momento que le precede, o la regresión causal derivada de las proposiciones de que todo suceso tiene una causa separada y que cada causa es un suceso. Podríamos incluso aceptar la regresión causada por la sugerencia de que, cuando creemos que *p*, creemos que *p* es probable (la regresión se da al considerar *q* = «*p* es probable»). ¿No pode-

mos aceptar simplemente que la justificación continúa *ad infinitum*? Creo que el regreso de justificaciones, una vez que se le permite comenzar del modo que se ha indicado, es vicioso, en el sentido de que mostrará que ninguna creencia está nunca justificada. Sin embargo, hay una razón para pensar tal cosa que es incorrecta: considerar que la regresión es temporal, considerar que *antes* de justificar A se deben justificar B y C, y así *ad infinitum*, de modo que nunca se podría comenzar. No considero que el argumento de la regresión esté preocupado con relaciones temporales entre actos de justificación. Un punto de vista mejor se limita a subrayar lo que dijimos anteriormente, que la regresión muestra que, si toda justificación es inferencial, ninguna creencia está justificada más que condicionalmente. Si el conocimiento requiere algo más que justificación condicional, como parece, la única manera de escapar al escepticismo del argumento de la regresión es concluir con el fundamentalista que hay creencias justificadas de un modo no inferencial.

Para evitar esta conclusión fundamentalista, deberemos mostrar que el argumento del regreso de justificaciones es falaz. Más adelante (9.1) se le dará una respuesta no fundamentalista. Por el momento, sólo deseo dirigir la atención a la posible ambigüedad de uno de sus pasos cruciales. La oración «Sólo hemos mostrado que A está justificada si lo están B y C» podría querer decir, como supusimos anteriormente, que habíamos mostrado que la justificación de A estaba condicionada a la justificación de B y de C; pero también podría querer decir que, cuando B y C están justificadas de hecho, hemos mostrado que lo está A, es decir, que lo que es condicional no es la justificación que hemos demostrado, sino el éxito de la propia demostración. El argumento presentado anteriormente requería la primera interpretación de esa oración crucial. Si la interpretamos de la otra manera, lo que tenemos no es una regresión de las justificaciones, sino una demostración de la justificación que sólo tiene éxito en ciertas condiciones.

El argumento del regreso de justificaciones se diferencia del argumento de Lewis sobre probabilidad y certeza, a pesar de sus grandes semejanzas estructurales (de hecho, ambos son argumentos sobre regresiones). La diferencia entre los dos radica en que la primera de las regresiones sólo puede ser detenida por ciertas creencias (infalibles), mientras que el segundo insiste tan sólo en la existencia de creencias que no están justificadas inferencialmente.

INFALIBILIDAD Y JUSTIFICACIÓN

La idea de que cualquier creencia infalible habrá de justificarse de un modo no inferencial unifica los dos argumentos previos, como un argumento conjunto a favor del fundamentalismo clásico. Una creencia infalible estaría justificada, sin derivar su justificación de ninguna relación con otras creencias; no necesitaría ningún tipo de apoyo externo, dado que parece obvio el carácter impecable de cualquier creencia cuyas posibilidades de ser falsa son nulas. Por tanto, si hay creencias infalibles, no debemos preocuparnos por una regresión amenazante de justificaciones. La infalibilidad de la base detendrá tal regresión.

Veremos en 4.3 que, aunque todas las creencias infalibles estén justificadas de un modo no inferencial, lo contrario no es cierto. Es esto lo que abre la puerta a tipos de fundamentalismo distintos a los de la versión clásica. Es posible abandonar la opinión de que tenemos creencias infalibles y encontrar maneras diferentes de suponer que algunas creencias están justificadas de un modo no inferencial. Pero no podemos limitarnos a presumir que tal cosa ha de suceder, por ejemplo, con nuestras creencias sobre estados sensoriales. Debemos elaborar alguna explicación de cómo es posible que una creencia logre *status* y desempeñe ese papel especial. El fundamentalismo clásico dice, plausiblemente, que nuestras creencias sobre nuestros estados sensoriales lo consiguen porque son infalibles. En la sección siguiente argumentaré que tal cosa no puede ser cierta y, si mi argumento es correcto, deberemos encontrar alguna otra manera de mostrar cómo nuestras creencias pueden estar justificadas de un modo no inferencial y sustentarse, de ese modo, sobre sí mismas.

4.2. ALGUNOS PROBLEMAS PARA EL FUNDAMENTALISMO CLÁSICO

Una de las razones más importantes para desear que las propias creencias clásicas fueran infalibles es que tal cosa garantizaría que todas fueran verdaderas. ¿Tiene algún propósito la búsqueda de esta garantía? Los principios de inferencia, por los que nos movemos de creencias básicas a creencias no básicas, son falibles, en el sentido de que, a veces, nos llevan de creencias verdaderas a falsas. (Recordemos el pollo de Russell [Russell (1959), p. 35], cuyas creencias verdaderas sobre la regularidad con la que había sido alimentado hasta un

momento determinado le llevaron a una falsa creencia sobre la seguridad de su futuro.) Si el procedimiento incorpora siempre esta fuente de contaminación, ¿qué sentido tiene insistir en que el *input* esté completamente esterilizado, limpio de cualquier tinte de falsedad?

Pero la objeción principal al fundamentalismo clásico es la de que no existen creencias infalibles. El *falibilista* sostiene, en mi opinión correctamente, que en ningún lugar estamos completamente a salvo de la posibilidad del error.

¿Son infalibles las creencias sobre nuestros estados sensoriales presentes? Los defensores más destacados de la infalibilidad tienden a conceder que hay espacio para el error en la *descripción* de los propios estados sensoriales [Ayer (1950)]. Podría equivocarme al describir mi experiencia como de rosa cuando en realidad fuera de naranja. Pero se supone que sólo sería un error verbal. Por supuesto, puedo estar equivocado sobre los significados de las palabras que utilizo, pero tal cosa no mostraría que me equivoco respecto a mis estados sensoriales presentes. Por el contrario, debo saber cómo me parecen ser las cosas; mi único error radica en escoger las palabras erróneas. La descripción que uso podría ser falsa, pero yo continuaría siendo infalible. *Mis creencias* —las cosas que mis palabras tratan de describir con más o menos éxito— deben ser verdaderas.

De un modo semejante, podemos decir que los errores meramente verbales pueden corregirse por los métodos habituales. Puedes mostrarme o recordarme las diferencias entre el rosa y el naranja, quizá enseñándome una carta de colores. Cuando haya comprendido la diferencia, la podré aplicar a mi experiencia presente para poder decidir si es de naranja o de rosa. Pero para ello debo ser previamente consciente de la naturaleza de la experiencia, antes de poder compararla con otras para decidir las palabras que la describen correctamente. No cambio de opinión sobre cómo me parecen las cosas, sólo respecto a cómo describirlas.

En tercer lugar, aunque yo necesite comparar entre mi experiencia presente y otras para que saber qué palabras utilizar en las descripciones, y aunque semejante comparación, especialmente en el caso en que comparo una experiencia presente y una pasada, sea falible (dado que la memoria es falible), no es la comparación lo que trato de expresar cuando trato de expresar mis creencias sobre mi mera experiencia presente. Dado que mi experiencia hubiera sido como es, independientemente de otras experiencias que hubiera podido tener o no tener. De modo que la falibilidad de la comparación no llega a mostrar la falibilidad de la expresión de creencia.

Por último, si la comparación es de alguna forma posible, sólo puede ser porque tenemos conocimiento no–comparativo de las dos cosas comparadas. Las comparamos para ver no cómo es cada una, sino en qué aspecto se parecen.

¿Qué replica debe dar el falibilista a este argumento? En primer lugar, ¿cuál es el contenido de semejante creencia infalible? No puede ser que la manera en que las cosas me parecen ahora a mí es de color rosa, dado que podría dudar de si semejante manera de parecer es o no rosa. Es más verosímil que la creencia infalible lo sea respecto al hecho de que las cosas *me parecen ahora de esa manera*. Pero ¿a qué equivale semejante creencia? ¿Qué contenidos tiene? No es una descripción interna y, de algún modo, no verbal de cómo parecen las cosas; no dice cómo parecen. A lo único a lo que equivale semejante creencia es a un gesto hacia algo; en realidad, una especie bien extraña de gesticulación, dado que los gestos son normalmente inteligibles sólo como actos públicos respecto a objetos públicamente observables, mientras que aquí nos las vemos con un acto privado de gesticulación con un objeto privado. Si los gestos dirigen la propia atención hacia algo, ¿es inteligible la idea de nosotros mismos dirigiendo nuestra atención a cómo nos parecen las cosas?

El infalibilista puede insistir en que no puedo equivocarme al creer que el color rosa es la manera en que me parecen las cosas, aunque pueda equivocarme al suponer que «rosa» es la palabra que debe usarse para describir la manera en que las cosas me parecen. Éste es el movimiento del «error meramente verbal». Pero se diría que esto es un mal uso de la noción de un error *meramente* verbal. Hay bastantes clases de errores verbales (un estudio de la vida de Warden Spooner sería instructivo en este punto, aunque no pueda evitarse la sospecha de que algunos de sus errores fueron deliberados). Pero el caso en el que, escogiendo las palabras cuidadosamente y con plena conciencia de lo que estoy haciendo, me pronuncio deliberadamente sobre la naturaleza de mi estado sensorial presente no es uno de ellos. En este caso, si estoy equivocado, mi error es sustancial, dado que, al estar equivocado respecto a que «rosa» sea la palabra adecuada para describir mis experiencias presentes, estoy equivocado sobre qué sea el color rosa y, por ello, sobre si mi experiencia es de rosa más bien que de naranja. De modo que, en este caso, el error es tanto verbal como sustancial.

Si el contenido de una creencia que se supone infalible es simplemente que las cosas me parecen de cierta manera ahora, hay obviamente menos espacio para el error que si me arriesgara a creer que esa manera particular es de color rosa. A menos contenido, menor riesgo

y mayor posibilidad de infalibilidad. Por lo tanto, parece probable que una creencia sólo puede ser realmente infalible si está privada de todo contenido. Esta es la conclusión falibilista extrema. Pero, incluso si esta conclusión no está justificada, podemos decir que las creencias infalibles deben tener un contenido mínimo y evanescente. Lo importante de ello es que el programa del fundamentalismo clásico pretendía que las creencias infalibles fueran aquellas por referencia a las cuales las demás creencias se justificaran. Se suponía que eran las creencias *básicas* que fundamentan a las demás, los auténticos fundamentos epistemológicos. Para cumplir tal papel necesitarían tener el contenido mínimo que les permitiera funcionar como premisas de inferencias. Con la reducción de contenido necesaria para mantener la infalibilidad de esas creencias básicas parece improbable que las creencias interesantes sobre el pasado, el futuro, lo inobservado o, incluso, los contextos materiales presentes pudieran justificarse apelando a ellas. Nuestras creencias básicas deben tener suficiente contenido como para apoyar la superestructura en la que estamos interesados, y ninguna creencia con semejante grado de contenido podrá ser infalible.

Una breve consideración de los argumentos de Chisholm, destacado defensor del fundamentalismo contemporáneo, puede confirmar este diagnóstico de los errores del infalibilismo. Chisholm distingue entre el uso comparativo y el no comparativo de la frase «parece blanco» (1977, pp. 30-33). En el uso comparativo, «X parece blanco» es una abreviación de «X parece como que parecen normalmente las cosas que son blancas». Pero el uso no comparativo, que se da en la oración «las cosas blancas normalmente parecen blancas», es muy diferente. La última oración se convertiría en una tautología si la analizáramos como si involucrara el uso comparativo. No se trata de una tautología, por lo que debe haber un uso distinto, no–comparativo, de la expresión «parece blanco»: un uso en el que realizamos un genuino intento de describir, sin comparar, el modo como parecen normalmente las cosas blancas. Chisholm pretende que en el uso no–comparativo los enunciados sobre apariencias expresan lo que es «directamente evidente». Una proposición directamente evidente es una que, en terminología de Chisholm, no es ni idéntica a, ni está implicada por, una proposición contingente verdadera que no llega a ser cierta. (Una proposición contingente es aquella que podría ser o no verdadera o falsa.) Una creencia en una proposición directamente verdadera no es lo mismo que una creencia infalible, aunque las dos comparten la característica que nos importa aquí: ambas son verdaderas.

Chisholm considera diversas objeciones a su tesis de que hay un

uso no comparativo de «parece blanco» que expresa lo que es directamente evidente (y, por tanto, verdadero). Algunas de sus observaciones se han hecho eco de los argumentos a favor del infalibilismo que se han dado anteriormente. La última objeción que considera reza del siguiente modo (1977, p. 33):

> *a)* al decir «Algo parece blanco» hacemos ciertas asunciones sobre el lenguaje; suponemos, por ejemplo, que la palabra «blanco» y la frase «parece blanco» se usan de la misma manera en la que las hemos usado en otras ocasiones, o en la que otras personas las han usado. Por tanto, *b)* cuando decimos «esto parece blanco», estamos diciendo algo no sólo sobre la experiencia presente, sino también sobre todas esas ocasiones. Pero *c)* lo que estamos diciendo sobre todas esas ocasiones no es directamente evidente. Y, por tanto, *d)* «esto es [*sic*; debe ser «parece» (J. D.)] blanco» no expresa lo que es directamente evidente.

Chisholm comenta correctamente que en ese argumento el error está en el tránsito de *a)* a *b)*. Estoy de acuerdo en que no es un paso válido, *b)* debería ser *b'):* «Cuando se dice "Esto parece blanco", lo que se dice no puede ser verdadero a menos que ciertas proposiciones sobre experiencias de otros distintos al que habla sean verdaderas.» Todo esto genera un argumento mucho más fuerte, pero lo que me interesa en este punto es la razón de Chisholm para mantener que *b)* no se sigue a *a)*. Afirma:

> Debemos distinguir entre la creencia que un hablante tiene sobre las palabras que está utilizando y la creencia que trata de expresar por medio de esas palabras. Un francés [puede creer que] «patatas» en castellano significa «manzanas»...; del hecho de que tenga creencias equivocadas sobre las palabras «patatas» y «manzanas» no se sigue que tenga creencias equivocadas sobre las patatas y las manzanas.

Es obvio de Chisholm sigue en este punto la línea tradicional de que todo error aparente respecto a las creencias sobre nuestros estados sensoriales es un error verbal, y que el error verbal debe distinguirse del error sustancial. Esto confirma nuestro diagnóstico original sobre el paso erróneo que da el infalibilista. Chisholm da ese paso, y no proporciona ninguna razón nueva para creer en la infalibilidad.

4.3. FUNDAMENTALÍSIMO SIN INFALIBILIDAD

En ausencia de la infalibilidad, el programa del fundamentalismo clásico se viene abajo, pero no vimos ninguna razón para suponer que

ésta era la única, ni siquiera la mejor, forma de fundamentalismo. ¿Qué versiones más débiles son posibles?

La primera tesis característica del fundamentalismo involucra una respuesta al argumento del regreso de justificaciones:

> F^1: Hay dos formas de justificación, la inferencial y la no–inferencial.

Pero la aceptación del argumento del regreso de premisas involucra una tesis adicional:

> F^2: Las creencias básicas nunca están justificadas, ni siquiera en parte, por apelación a las creencias no–básicas.

F^2 es incompatible con cualquier sugerencia de que la justificación no–inferencial de las creencias básicas es sólo parcial y necesita ser complementada por otras creencias. Esta sugerencia podría tentar a aquellos que advierten que, por más que normalmente aceptemos sin ninguna reticencia la descripción que alguien da de sus propios estados sensoriales, a veces la objetamos diciendo, por ejemplo: «No es posible que ese semáforo te parezca naranja. Está encendida la luz de arriba, que es siempre la roja.» Si las creencias no básicas de este tipo pueden reducir la justificación de las creencias básicas, es de presumir que también pueden incrementarla, en cuyo caso se daría la posibilidad de que, aunque las creencias sobre nuestros propios estados sensoriales estén siempre parcialmente justificadas precisamente por su propio objeto, no estuvieran justificadas, completa y satisfactoriamente, a menos que se diera evidencia confirmatoria, o que no hubiera evidencia refutante, al nivel no–básico. Pero semejante idea no sería accesible a alguien convencido por el argumento del regreso de justificaciones. El fundamentalismo supone que hay dos tipos de justificación y que las creencias justificadas inferencialmente se justifican por apelación a las que están justificadas de un modo no inferencial. Si continuamos admitiendo que las últimas se justifican parcialmente por apelación a las primeras, reintroducimos el círculo de la justificación condicional cuya consecuencia escéptica era la de que nada estaba realmente justificado. El fundamentalismo debe mantener la dirección de la justificación en un sólo sentido, desde el no–básico al básico, o renunciar a cualquier posible utilización del argumento del regreso de justificaciones.

Así pues, el único modo de fundamentalismo que podría retener F^1

sin F^2 sería el que aceptara F^1 por razones distintas de aquellas que proporcionaba el argumento del regreso de justificaciones. Podría sostenerse que las creencias sobre nuestros propios estados mentales están justificadas siempre hasta algún grado, precisamente por su objeto (por ende, no inferencialmente), mientras que el resto de las creencias, si están justificadas, tienen una justificación inferencial. Podría suponerse que todo esto es un intento de dar sentido a la idea empirista de que nuestras creencias sobre la experiencia presente tienen una estabilidad de la que carecen otras creencias, en virtud de lo cual son capaces de justificar esas otras creencias y adecuarse así a la exigencia empirista de que (por expresarlo vagamente) todo el conocimiento esté basado en la propia experiencia. Este nuevo tipo de fundamentalismo, que se considerará más adelante en el capítulo 6, podría escapar a las exigencias de F^2, pero sólo a costa de abandonar el arma favorita del fundamentalismo, el argumento del regreso de justificaciones.

Cualquier fundamentalista tiene el deber de dar sentido a la posibilidad de que haya creencias justificadas no inferencialmente. ¿Cómo podrían ser? El requisito formal que impone el argumento del regreso se vería satisfecho si hubiera creencias de uno cualquiera de los tres tipos siguientes:

1. creencias que están justificadas por algo distinto a creencias;
2. creencias que se justifican a sí mismas;
3. creencias que no necesitan justificación.

Sería poco apropiado eliminar 3 por la mera razón de que el argumento pide que sólo las creencias justificadas puedan justificar otras; si podemos dar un sentido razonable a la noción de una creencia sustentándose sobre sí misma, más bien que sobre otras, sería fácil considerar 3 como un caso especial de 2.

Debemos considerar 1–3 como propiedades formales; las creencias con esas propiedades detendrían la regresión, pero cualquier propiedad formal como ésas necesita fundamentarse en una propiedad «epistémica» más fundamental. La infalibilidad es una propiedad de este tipo, como se puso de relieve anteriormente. El fundamentalismo clásico suponía quizá que las creencias infalibles serían del tipo 2. En ausencia de infalibilidad, ¿qué movimientos similares nos son accesibles?

C. I. Lewis afirmaba que las creencias básicas eran «ciertas» o «incorregibles» [Lewis (1952 y 1946), cap. 7]; no está siempre claro si pensaba de ellas que eran infalibles. De un modo similar, otros (Descartes, quizá) sostuvieron que eran o podían ser «indudables».

Podríamos definir la incorregibilidad e indubitabilidad del siguiente modo:

> Una creencia es *incorregible* si y sólo si nadie podría estar en posición de corregirla.
>
> Una creencia es *indudable* si y sólo si nadie podría tener nunca una razón para dudarla.

Cualquiera de estas dos propiedades ¿es capaz de proporcionarnos una forma clásica de fundamentalismo que, aunque débil, sea todavía atractiva? Según mi punto de vista, una vez que admitamos que nuestras creencias sobre nuestros estados sensoriales no son infalibles y pueden ser falsas, la incorregibilidad sería un vicio más que una virtud. El pensamiento de que algunas creencias básicas son incorregiblemente falsas es demasiado terrible como para tenerlo en cuenta. De igual modo, podríamos preguntarnos cómo podría ser indudable una creencia falible, dado que las creencias básicas están en relaciones inferenciales con creencias mucho más interesantes sobre los objetos físicos, y tales creencias son ciertamente dubitables; siempre existe la posibilidad de que alguien tenga una razón para dudarlas. Pero, si es así, es difícil ver cómo la dubitabilidad de las creencias no básicas no se transmite a las básicas sobre las que descansan; parece obvio que la falsedad de una creencia no básica sería una razón para dudar la creencia básica sobre la que descansa la primera, una vez admitido el hecho de que las creencias básicas *pueden* ser falsas.

De modo que ni la incorregibilidad ni la indubitabilidad pueden proporcionarnos una forma alternativa de fundamentalismo. Pero hay otras formas posibles. Podríamos sugerir la posibilidad de que hubiera creencias del tipo 1 si hubiera creencias que se justificaran por apelación a los hechos, y que una creencia tal estuviera justificada si la causaran los hechos. Austin propuso la idea de que, en la ocasión apropiada, lo que justifica mi creencia de que hay un cerdo delante de mí es, simplemente, el cerdo [Austin (1962), pp. 115-116]. Pero esta idea no se puede generalizar fácilmente a no ser que supongamos que la justificación se alcanza menos por apelación al cerdo que por apelación al hecho de que hay un cerdo delante de mí; el hecho, en este caso, causa la creencia. Este movimiento debe recordarnos algunas observaciones hechas en 2.4 sobre la teoría causal del conocimiento. Aceptamos la idea de que los hechos pueden ser causas; la dificultad ya mencionada de que los hechos universales no pueden causar creencias universales no importa ahora, dado que no es probable que las creencias universales sean básicas.

Otra posible versión de fundamentalismo sostiene que hay algunas creencias que nos están dadas como «datos», y que están completamente justificadas a menos que surja algo que refute su justificación (cf. el uso de la refutabilidad en 2.3). Podríamos denominar esta justificación «*prima facie*» o «refutable»; es más débil que la que nos da la indubitabilidad, dado que contempla la posibilidad de que haya razones en contra de una creencia básica. Pero todavía acepta tanto F^1 como F^2.

Ya ha sido propuesta una versión final. Más débil que la última, sostiene que las creencias dadas como «datos» no están nunca justificadas completamente por esa mera razón, pero que tales creencias ya están justificadas parcialmente, independientemente del apoyo que puedan recibir de otras creencias. Sin embargo, sin ese apoyo ulterior su justificación es insuficiente. Éste es el fundamentalismo que acepta F^1 y no acepta F^2.

Estas versiones diferentes del fundamentalismo no se ven afectadas por la falta de infalibilidad. En el próximo capítulo consideraremos un problema diferente para el fundamentalismo que pueda tener efectos desastrosos.

LECTURAS COMPLEMENTARIAS

C. I. Lewis (1952) y Ayer (1950) defienden formas de infalibilismo.

Alston (1976) y Amstrong (1973, cap. 11) ofrecen un buen análisis del argumento del regreso.

Firth (1964) clasifica las formas diferentes de fundamentalismo desde el punto de vista de su postura rival, el coherentismo.

La noción de infalibilidad que se usa en este capítulo precisa de detallada atención. Alston (1971) analiza las diferencias entre infalibilidad, incorregibilidad, indubitabilidad, etc.

C. I. Lewis (1952) es una respuesta a Reichenbach (1952) y Goodman (1952), que argumentan contra su opinión de que nada es ni siquiera probable si no hay nada que sea cierto.

La mayor parte de los textos anteriores, con otros muchos ensayos importantes sobre temas tratados en este capítulo, están reimpresos en Chisholm y Swartz (1973).

El fundamentalismo de Chisholm (1977) descansa explícitamente en el argumento del regreso en las pp. 16-20.

Todavía vale la pena leer el ataque de Austin (1962) a Ayer, especialmente el capítulo 10.

Sellars (1963, cap. 5) constituye un ataque extremadamente influyente al «mito de lo dado» que se incorpora al fundamentalismo clásico. Se trata de un ensayo difícil que recompensa su estudio.

5. EL FUNDAMENTALISMO Y LAS OTRAS MENTES

5.1. LAS CREENCIAS BÁSICAS Y LOS PROPIOS ESTADOS SENSORIALES

En el último capítulo se presentó el fundamentalismo clásico como una respuesta a tres tipos de exigencias: las de dos argumentos sobre regresiones y las del empirismo. La doble respuesta consiste en suponer que hay creencias básicas infalibles y que esas creencias tratan de la naturaleza de los propios estados sensoriales. En 4.2 se dieron razones para dudar de la existencia, o la utilidad, de las creencias infalibles. Me gustaría examinar ahora la otra mitad de la historia.

Hay una fuerte tradición en filosofía que mantiene que nuestro punto de partida es el conocimiento de nuestros estados sensoriales y que todo lo demás se construye a partir de ahí. Descartes nunca cuestionó sus creencias sobre cómo le parecía que eran las cosas en un momento dado; se preguntó, más bien, cómo podría saber otras cosas, tales como la existencia de Dios o del mundo material. Es posible que el patrón clásico sea el que estableciera John Locke, al mantener que una persona es directamente consciente sólo de la naturaleza de sus propias ideas y que todo lo demás, si se conoce, se conoce de un modo indirecto. Este punto de vista ha sido bastante habitual en epistemología desde entonces. Es posible que conozcamos la naturaleza de nuestros propios estados sensoriales, pero ¿cómo construir a partir de ahí para llegar al conocimiento del pasado, del futuro o de los estados sensoriales de los demás?

Un aspecto cuestionable de este punto de vista, que parece auténticamente cartesiano, es la sugerencia de que la epistemología se centra en el individuo. No tiene ningún interés en poner de relieve el crecimiento del conocimiento a lo largo de las generaciones, o la fecundidad de la herencia que nos legaron los investigadores que nos han precedido. Más bien, se piensa en cada uno de nosotros como si comenzara más o menos de nuevo, y la cuestión filosófica es la de explicar

cómo podemos llegar a partir de tal estado, por medio de nuestra conciencia de la representación cambiante de la experiencia sensorial, al sofisticado conocimiento que todos tenemos como miembros de una sociedad moderna.

En este capítulo daré diversas razones para rechazar ese punto de vista, o, si no para rechazarlo, para mostrar que las cuestiones que plantea nunca pueden encontrar respuesta, de modo que conduce directamente al escepticismo. De hecho, hay una tendencia escéptica en este tipo de fundamentalismo, precisamente porque nos lleva a considerar problemático todo lo que no sea el conocimiento de nuestros propios estados sensoriales; reconoce el peligro de que pudiéramos ser incapaces de construir el tipo de superestructura que se supone sustentada por tales cimientos.

5.2. EL PROBLEMA DE LAS OTRAS MENTES

Una de esas áreas en las que existe el peligro de que no sepamos nada es la de las otras mentes. Cada uno de nosotros conoce la naturaleza de sus propios estados sensoriales, pero ¿podemos conocer la naturaleza de los estados sensoriales de los demás? Aún más, ¿podemos saber si hay otras mentes que tienen estados sensoriales distintos a los de cada uno de nosotros?

El mero hecho de denominar el problema como «Problema de las Otras Mentes» consagra el fundamentalismo. El conocimiento de nosotros mismos es seguro, y el de los otros es problemático. (Del mismo modo, el apelativo tradicional de otro problema, «Nuestro Conocimiento del Mundo Externo», consagra cierto tipo de aproximación epistemológica sin la que no habría problema, o, al menos, *ese* problema. Pudiera ser que tal aproximación fuera la que, más que solucionar el problema, lo originara.) A pesar de ello, aceptaremos esa perspectiva por el momento. No será por mucho tiempo.

No hay ninguna duda de que, de acuerdo con el punto de vista tradicional, tendremos dificultades para afrontar los argumentos en favor del escepticismo sobre las otras mentes. Puede usarse el argumento del error para mostrar que sabemos de casos en los que los demás han conseguido ocultar sus propios estados mentales, o han fingido estar en un estado mental distinto al real (en realidad, no se trata de alternativas distintas). ¿Qué nos asegura que un nuevo caso no es de esos tipos?

No hay nada de incorrecto en esa formulación del argumento, pero es débil. Da por sentado que hay otras mentes y se limita a argüir que

podemos estar equivocados respecto a sus estados, pensando que alguien es feliz cuando está triste, etc. Para usar el argumento del error sin semejante supuesto, deberíamos ser capaces de señalar casos en los que nos equivocamos al considerar que se daba una mente. Es posible que existan tales casos, gracias al creciente desarrollo de la tecnología, con máquinas que juegan muy bien al ajedrez, o que son capaces de contestar coherentemente.

Pero lo que se usa para generar el escepticismo sobre la existencia de mentes distintas a la propia es, normalmente, la apelación a casos posibles más bien que reales. ¿No podría ser el caso que esos otros cuerpos con los que parece que tenemos relaciones personales de amistad, o enemistad, carecieran en realidad de cualquier estado mental, que sólo tuvieran conducta exterior desprovista de toda vida interior? Ninguno de nosotros podría distinguir del mundo real un mundo en el que la conducta de los demás fuera la misma, pero en el que no tuvieran vida mental. ¿Cómo sabemos, entonces, que en realidad hay otras mentes? Por todo lo que podemos distinguir, no hay ninguna.

5.3. EL ARGUMENTO POR ANALOGÍA

El argumento por analogía que vamos a tratar admite que *es posible* que los objetos que denominamos personas, distintos a cada uno de nosotros, sean autómatas sin mente, pero afirma que, a pesar de ello, tenemos suficientes razones para creer que ése no es el caso. Hay más evidencia a favor de que no son autómatas sin mente que a favor de que lo sean.

La formulación clásica del argumento se encuentra en J. S. Mill. Escribe (1867, pp. 237-238):

> Soy consciente en mí mismo de una serie de hechos conectados por una secuencia uniforme, el principio de la cual es la modificación de mi cuerpo, el punto medio son las sensaciones, y el final, la conducta exterior. En el caso de los otros seres humanos, tengo la evidencia de mis sentidos para el primer y último eslabón de esa serie, pero no para el eslabón intermedio. Sin embargo, descubro que la secuencia entre el primero y el último es tan regular y constante en esos casos como en el mío. En mi propio caso, sé que el primer eslabón produce el último a través del eslabón intermedio y que no lo podría producir sin él. Por tanto, la experiencia me obliga a concluir que debe existir un eslabón intermedio; que debe ser en otros, o bien el mismo que en mí, o bien otro diferente; [...] al suponer que el eslabón es de la misma naturaleza [...] actúo de acuerdo con las reglas legítimas de la investigación experimental.

La primera objeción podría consistir en decir que este argumento, como argumento inductivo, es muy débil, dado que está obligado a argumentar a partir de una única instancia. Pero esta crítica no da en lo esencial. Por la misma naturaleza del caso, la hipótesis de que las que parecen ser otras personas tienen sensaciones que son como las mías, aunque necesariamente inobservables, no es una hipótesis contra la que pudiéramos tener ninguna evidencia en contra. Si el argumento de Mill consigue mostrar que hay aunque sólo sea una pequeña evidencia a favor, ya habrá más evidencia a favor que en contra, por lo que deberemos aceptar la hipótesis.

Las dificultades del argumento derivan de dos de sus supuestos: *a)* sobre la separabilidad y *b)* sobre la comprensión. Acepta sin reticencia alguna *a)* la *separación de la mente (inobservable) respecto a la conducta* (observable) sobre la que se basa el argumento escéptico. El escéptico argumenta que el vínculo entre ambas cosas es meramente contingente, es decir, que es perfectamente posible que una se dé sin la otra, y que, por todo lo que sabemos, es eso lo que sucede. El argumento por analogía concede la primera parte y niega la segunda, afirmando que, como mínimo, tenemos evidencias de lo contrario. Al conceder la primera parte, muestra por qué es necesario el propio argumento por analogía. Es necesario porque no podemos considerar lo mental y lo físico como relacionados de un modo que no sea contingente. Si hubiera, por ejemplo, un vínculo *conceptual* entre estado mental y conducta tal que no pudiéramos concebir la presencia de la conducta sin el (o sin algún) estado mental, no sería preciso el argumento por analogía.

El argumento supone también *b)* que puedo *comprender qué es para los otros tener estados mentales*. Y proporciona una explicación rudimentaria de cómo puedo hacer tal cosa. Comprendo qué es el dolor a partir de mis propios recursos y del mejor de los modos posibles, es decir, sintiéndolo. Pero el concepto de dolor adquirido de ese modo puede aplicarse a los demás, por el mero supuesto de que ellos tienen una sensación similar a la que tengo yo (Mill: «suponiendo un eslabón de la misma naturaleza»).

La crítica principal al argumento por analogía es la de que esos dos supuestos son incoherentes. Esta crítica se la debemos a Wittgenstein y la expone Malcolm (1958). Si no hay un vínculo no–contingente entre conducta y estado mental [es decir, si aceptamos *a)*], nunca podremos mostrar que *b)* es verdadera, por lo que estaremos a merced de una forma excepcionalmente poderosa de escepticismo local (cf. 1.1). De modo que, si el argumento por analogía es necesario, por

a), está condenado al fracaso con respecto a *b)*. Por otra parte, si conseguimos mostrar *b)*, sólo lo podemos hacer negando *a)* y convirtiendo el argumento por analogía en algo completamente innecesario.

Todo esto no es más que una declaración de intenciones hasta que mostremos cuál es el argumento escéptico crucial. Debe ser distinto a los que mencionamos anteriormente, dado que no muestra razón alguna para dudar de nuestra comprensión de las proposiciones sobre otras mentes. Hasta ahora hemos dado por supuesto que comprendemos la sugerencia de que hay mentes distintas a la nuestra propia, y nos hemos limitado a preguntar cuánta evidencia tenemos a su favor. El argumento que vamos a considerar ahora trata de mostrar que este supuesto (característico del escepticismo débil) es falso.

5.4. ¿PODEMOS COMPRENDER PROPOSICIONES SOBRE MENTES DIFERENTES A LA NUESTRA?

¿Por qué la separación de lo físico y lo mental convierte en imposible comprender qué sea la existencia de mentes distintas a la nuestra? El argumento por analogía supone que podemos construir, a partir del caso propio, el concepto de dolores que pueden sentir los demás y no cada uno de nosotros. Pero ¿es tan fácil concebir un dolor que no me duela a mí? Deberé empezar con uno de mis dolores y hacerme a la idea de que existe algo similar que duele, pero sin que me duela a mí, y también a la idea de que podría haber algo que fuera similar a mí, sin ser yo, al que los dolores le duelen. Ambas ideas son bastante dudosas. ¿Cómo concebir algo doloroso sin concebir que me duele a mí? Parece obvio que, al concebir algo que duela, necesariamente lo concibo como algo que *me* duele. Mi concepción de dolores que no son míos no es la concepción de algo que es como mi dolor, dado que duele, pero que no es de los míos. Del mismo modo, ¿me es realmente fácil concebir otro sujeto de experiencia distinto a mí? Si no lo puedo hacer a partir de la idea de dolores que no sean míos, no lo puedo hacer de ninguna otra manera. Es imposible sustraer de la idea de mi dolor aquella parte que es mi yo, el sujeto, y sustituirla por la noción de otro sujeto (algo similar a mí que no sea yo) y suponer que el resto (el dolor) permanece inalterado. Cualquier dolor que pueda concebir doliendo exactamente como me duele a mí, debe concebirse como doliéndome a mí.

Este argumento escéptico es crucial. Muestra que no tiene ningún sentido limitarse a proclamar la posibilidad de concebir el dolor de

otro sobre la única base de que es posible suponer que él tiene lo mismo que yo. Esa observación [véase *b)* anteriormente] sólo sienta la cuestión por anticipado en contra del argumento relevante.

Tampoco tiene sentido alguno insistir en que podemos concebir qué es que una rodilla distinta a la nuestra duela. Un dolor concebido como estando en otro cuerpo no es un dolor concebido como doliendo a otra persona. De hecho, si el argumento anterior es correcto, concebir una rodilla distinta a la mía y endolorecida como la mía es exactamente concebir que a mí me duele esa rodilla. Podríamos admitir, pues, que es posible concebir que tengo dolor en la rodilla de otro (un sentido nuevo de «Me duele su rodilla») pero, en este caso, no podemos concebir que la rodilla duela a alguien más.

Por tanto, el argumento escéptico pretende que yo no puedo dotar de sentido a la idea de un sujeto de experiencias distinto a mí mismo. Me es imposible concebir experiencias que no sean mías (generalizando a partir del ejemplo más simple del dolor) y tampoco tengo ninguna otra manera de concebir un sujeto de experiencias que no sea yo mismo. Nos conduce directamente desde el fundamentalismo (ya que el argumento se formuló en términos fundamentalistas) a la forma más interesante de *solipsismo*, el punto de vista de que debo considerarme a mí mismo como el único sujeto de experiencia —de hecho, como *el* sujeto de experiencia— ya que me es completamente imposible formarme concepción alguna de ningún otro sujeto.

¿Por qué muestra todo esto que el argumento por analogía, al aceptar el supuesto *a)*, no puede dar cuenta de *b)*, no puede explicar nuestra comprensión de los enunciados sobre otros sujetos? La razón es que muestra cómo, si nuestro único punto de partida es el caso propio y nos limitamos por entero a una concepción de la mente que convierte los estados mentales en algo independiente de la conducta, no podemos pasar de nuestra concepción de nosotros mismos como sujetos de experiencia a la concepción de otros sujetos. El concepto de estado mental incorporado al argumento, y sobre el que descansa el argumento escéptico, es demasiado restrictivo para que sea posible tal movimiento.

La conclusión más adecuada de todo esto no es que el escepticismo o el solipsismo sean inevitables, sino que los términos en los que se presentó el argumento escéptico, el supuesto *a)*, deben rechazarse. Y, al rechazar ese supuesto, hacemos que el argumento por analogía se convierta en un argumento redundante.

¿Cómo encontrar una concepción de los estados mentales distinta a aquella que maneja el argumento escéptico? Una de estas concepcio-

nés sería el *conductismo*, que proporciona un vínculo conceptual entre conducta y estados mentales, diciendo que mi estado mental cuando me duele la rodilla es exactamente el estado conductual en el que me inclino y me cojo la rodilla con sumo cuidado, etc. Hay también formas más débiles de conductismo que intentan contemplar la posibilidad de estados mentales ocultos; se podría mantener, por ejemplo, que el que yo tenga dolor en la rodilla es mi *disposición* a actuar de cierto modo, lo haga o no realmente de ese modo. Para ocultar mi dolor, deberé reprimir mi deseo (disposición) de inclinarme y de cogerme la rodilla con sumo cuidado, etc.

La dificultad obvia con las explicaciones conductistas de los estados mentales es la de que, aunque nos proporcionan de algún modo lo que estamos buscando, es decir, un análisis de los estados mentales en el que éstos están vinculados a la conducta de un modo no contingente, parecen pasar por alto todo lo que es típicamente mental. En el caso del dolor, pasan por alto, por ejemplo, el modo en el que el dolor es *sentido*. Puede que esto no sea todo lo relevante en el concepto de dolor, pero, ciertamente, es parte de él. Incluso si fuéramos capaces de mostrar la imposibilidad de que existiera gente que sintiera como nosotros sin tener las disposiciones a la conducta que nosotros tenemos, o viceversa, la maniobra del conductista parece demasiado extremista. Lo que nos gustaría hallar es una posición intermedia entre el conductismo, que identifica los estados mentales con alguna función de la conducta, y el punto de vista común al escepticismo y al argumento por analogía, que los separa en demasía. Es difícil en filosofía hallar compromisos satisfactorios, porque tienden a compartir los defectos de ambos extremos y a carecer de sus virtudes. Esperemos, sin embargo, que sea posible hallar uno en este problema.

La crítica del argumento por analogía que se dio anteriormente se la debemos a Wittgenstein. Su conclusión es la de que nuestra única esperanza de mostrar cierta comprensión de las proposiciones sobre los estados mentales de otros es la de aceptar una relación no–contingente entre los estados mentales y la conducta, eliminando, de este modo, la posibilidad de que se separen demasiado. La manera en la que el propio Wittgenstein lo hizo puede expresarse por medio de la controvertida noción de *criterio*. He aquí un análisis provisional de esa noción:

> A es *criterio* de B si y sólo si la verdad de A o el hecho de que A suceda es necesariamente una buena —aunque refutable— evidencia para la verdad de B, y, en ausencia de indicaciones de lo con-

trario, evidencia suficiente. De modo que, en casos favorables, la verdad de A, o el hecho de que A suceda, justifican perfectamente la creencia en B, o la aseveración de B,

y

Saber tal cosa es parte del tener competencia en el concepto B; parte de lo que es conocer el significado de «B».

Esta noción de criterio podría usarse para mostrar que el argumento por analogía no es necesario. Ya que, usándolo, podemos decir que cualquiera que comprenda el concepto de dolor sabe que ciertos tipos de conducta son criterios de las adscripciones de dolor. No es posible, por tanto, que existan seres que se comporten como nosotros y que carezcan de ciertos estados mentales, como el dolor. En casos en los que no hay evidencia al contrario, tal conducta cuenta como evidencia concluyente de que quien actúa de ese modo tiene dolor. Por supuesto, sí es posible que alguien finja que tiene dolor o que actúa movido por el dolor. En muchos de esos casos, la evidencia creada por la satisfacción de los criterios se ve refutada por otra evidencia; lo vemos, por ejemplo, retorciéndose y cogiéndose el estómago, pero sabemos que ha realizado el mismo tipo de gesticulación en este momento de la representación teatral todas las noches de la semana, por lo que no estamos justificados para creer algo que sí estaría perfectamente justificado por evidencias similares en otro contexto. Pero el aspecto esencial del fingimiento en una actuación es el de que ésta es una actividad que también tiene sus propios criterios. Necesita de un trasfondo que podría no ser accesible. Si estoy en una batalla y mi amigo pierde media pierna izquierda por una explosión y se dobla en agonía, no hay ninguna posibilidad de que esté actuando o fingiendo.

De modo que la noción wittgensteiniana de criterio parece proporcionar el tipo de compromiso que buscábamos. Sin embargo, es sumamente idiosincrásica, dejando a parte dudas exegéticas sobre su exacta formulación. No *necesitamos* ahora seguir esta puerta que nos abre Wittgenstein, al menos no todavía. Todo lo que tenemos por el momento es que sólo podemos comprender el discurso sobre las otras mentes si hay algún tipo de vínculo no contingente entre estados mentales y conducta. Si no lo hubiera, parecería inevitable el compromiso con el solipsismo.

5.5. EL ARGUMENTO DEL LENGUAJE PRIVADO: SEGUIR UNA REGLA

El solipsismo es una posición inestable. De hecho, sufre de muchas incoherencias. Ésta es una de las conclusiones de una familia de argumentos que dio Wittgenstein y que se conocen en su conjunto como el «argumento del lenguaje privado». El solipsista está en la situación en la que pretende tener un lenguaje para describir sus experiencias presentes y pasadas, y, quizá, para especular sobre el futuro. Y éste es un lenguaje *privado*; nadie más podría aprenderlo porque las experiencias por apelación a las cuales obtienen significado sus términos son experiencias privadas del hablante. Wittgenstein argumenta que tal lenguaje privado es imposible. El solipsista no puede tener siquiera ese conocimiento limitado de sus propias experiencias.

La interpretación del argumento del lenguaje privado es muy discutible. Comenzaré con un análisis debido, principalmente, a Saúl Kripke (1981).

¿Cómo adquiere su lenguaje el solipsista? Pretende saber, a partir de su propio caso, qué sea, por ejemplo, el dolor. Experimenta una sensación de cierto tipo y decide aplicar el término «dolor» a todas las sensaciones que son como ella. Dado que recuerda exactamente cómo era la sensación original, ha elaborado un concepto de dolor del que se seguirá, para cada sensación nueva, si debe o no denominarla «dolor».

Lo esencial del argumento del lenguaje privado es que el solipsista ni siquiera puede encontrar un punto de partida en todo esto. Nada de lo que hiciera, concentrándose de determinada manera en la sensación original y emitiendo la palabra «dolor», conseguiría proporcionar un significado a esa palabra. Para que una palabra tenga significado es preciso que existan reglas para su uso: reglas en virtud de las cuales la aplicación de la palabra pueda contar como correcta en algunos casos e incorrecta en otros. Es en virtud de esas reglas como podemos dar sentido a la idea de que estamos objetivamente en lo correcto al denominar «dolor» a la nueva sensación. El solipsista podría aplicar mal la regla sólo si ésta hubiera sido fijada previamente: en ese caso, podría pensar que estaba siguiendo la regla cuando de hecho no lo hacia así; podría parecerle que la nueva sensación era semejante a la sensación original, cuando de hecho no se pareciera. Para que haya objetividad es preciso que «no todo lo que parezca correcto sea correcto». ¿Cómo es posible que concentrarse en la sensación original capacite al solipsista para crear una regla en virtud de la cual puede decirse que es objetivamente verdad que la nueva sensación es dolor? No es posible.

La interpretación de Kripke del argumento del lenguaje privado sostiene que la respuesta negativa a esta última cuestión es consecuencia de consideraciones generales sobre reglas y objetividad.

Alejémonos del caso especial del solipsismo y los términos de sensación, y consideremos un caso que puede parecer más simple (de hecho, es el caso en el que es más difícil dar plausibilidad a lo que decimos). El ejemplo es el de una regla matemática; la regla para +2. Suponemos que, si alguien sabe la regla para +2, sabe, para cada número n, qué número es $n + 2$. La regla crea una serie; 0, 2, 4, 6..., y, a cada paso de la serie, existe una respuesta objetivamente correcta respecto a cómo seguir.

La cuestión es la de qué puede haber en lo que aprendemos, cuando aprendemos la regla para $n + 2$, que convierta en algo objetivamente correcto la continuación 20.002, 20.004, 20.006, y en algo objetivamente incorrecto la continuación 20.004, 20.008, 20.012. ¿A qué podríamos señalar, cuando alguien escogiera esta última expansión, para mostrarle que estaba equivocado? Suponemos que la persona cree que la regla aprendida le exige esta expansión, más bien que la que nosotros escogeríamos (que, por supuesto, es la correcta). ¿Qué hace que ella esté equivocada y nosotros no?

No parece que haya nada en los pasos anteriores de la serie que pudiera apoyar una continuación más bien que la otra. Hay, de hecho, una fórmula que genera la continuación «errónea» a partir de esos mismos primeros pasos; dado que ambas series comienzan de la misma manera, no podemos apelar a la manera en que una comienza para justificar nuestra preferencia por una determinada manera de continuar.

Podríamos sugerir que, al concebir la regla para +2, concebimos de alguna manera la totalidad de los casos, y es esto lo que convierte a la interpretación «correcta» en correcta. Sin embargo, esta sugerencia no es convincente. Incluso si, por causalidad, fuera cierto que hubiéramos pensado exactamente en esos miembros de las series, no sería ese hecho la razón por la cual deberían ser considerados como expansiones correctas de la serie. Y, en todo caso, ¿qué decir de las otras continuaciones igualmente correctas que nunca nos pasaron por la imaginación hasta que tuvimos que escribirlas?

De un modo similar, podríamos decir que, al concebir la regla, nos limitamos a pensar que debíamos añadir 2 cada vez. Pero la sugerencia tampoco funciona, en parte porque no se nos dice qué debe contar como «añadir 2», y en parte porque da por sentado cómo interpretar ese «cada», sin tener en cuenta que podríamos preguntar sobre «cada» lo mismo que preguntamos sobre «+2», dado que siempre existe la

posibilidad de que alguien, al cabo de un rato, comience a usar «cada» del modo en que nosotros usamos «cada otro», pretendiendo que así se mantiene en lo correcto.

La última posibilidad parece ser la de que nuestra concepción original de la regla no era otra que la de la creación de una disposición a ejecutar la serie de un modo más bien que de otro. Pero ¿en virtud de qué hecho se creó una disposición para continuar *ahora* de un modo más bien que de otro? El contenido de mi disposición original parece determinado por el modo en que descubro en un momento dado mi tendencia a continuar, más bien que al revés.

Parece que hemos agotado todas las respuestas posibles a nuestra cuestión. ¿Deberemos concluir que el escepticismo aparentemente arbitrario que subyace a la cuestión no tiene, de hecho, ninguna respuesta, que no hay ninguna manera objetiva de seguir la regla? Esto parecería atentar contra nuestra intensa convicción de que 20.002, 20.004, 20.006 *es* la manera correcta de continuar y de que las otras alternativas son erróneas. Pero éste parece ser uno de los casos en los que el mero pensamiento no puede ser tan pretencioso. Afortunadamente, existe una posibilidad que hemos pasado por alto, dado que, hasta ahora, nos hemos concentrado exclusivamente en los recursos internos al individuo que sigue reglas, sobre el tipo de cosas a las que podría señalar el solipsista.

En consecuencia, Kripke ve que la conclusión del argumento es la de que, para encontrar una base de nuestra creencia de que existe un método objetivamente válido de continuar la serie, debemos buscar, más allá del individuo, en lo que sucede a la comunidad de seguidores de reglas. En último término, el fundamento de la objetividad no radica en el pasado, en la naturaleza de los primeros miembros de la serie o en nuestra anterior comprensión de la regla, sino en la conducta actual de nuestra comunidad lingüística o matemática. Lo que convierte en correcta nuestra continuación de la serie es el hecho de que la comunidad coincide con nosotros. Lo que convierte en incorrecta la expansión alternativa es el hecho de que sería única dado que, en este caso, «corrección» significa «actuar en coincidencia con los otros». Es fácil comprender la razón por la que la interpretación de Kripke se denomina «interpretación comunitaria».

Estas ideas sobre el seguimiento de reglas en matemáticas pueden generalizarse para que nos proporcionen el argumento sobre el solipsismo que estamos buscando. El solipsista no admite comunidad alguna para fundamentar su creencia de que es objetivamente verdadero que su sensación nueva es de dolor. Lo que quiere decir que, en un

caso nuevo, carece de toda regla que le lleve en una u otra dirección. Puede decir lo primero que se le ocurra con relación a cualquier regla, y será tan correcto como cualquier otra cosa. Lo que quiere decir que sus palabras pierden todo significado y se convierten en algo vacío. Si no importa lo que digamos, sería mejor quedarse callado. Por tanto, es imposible la existencia de un lenguaje privado como el que necesita el solipsista. Todo el lenguaje posible es necesariamente público, dado que no tiene significado alguno a menos que sea usado por una comunidad.

5.6. OTRA INTERPRETACIÓN

La interpretación comunitaria del argumento del lenguaje privado lo considera una consecuencia directa de las ideas de Wittgenstein sobre reglas, seguimiento de reglas y objetividad. Hacia el final de su discusión sobre seguir una regla, Wittgenstein dice (1953, § 202):

> Y *pensar* que uno está obedeciendo una regla no es obedecer una regla. De modo que no es posible obedecer una regla «privadamente»; de otro modo, pensar que uno estaba obedeciendo una regla sería lo mismo que obedecerla.

Y Kripke considera este párrafo como una breve anticipación del argumento del lenguaje privado, que oficialmente parece comenzar unas cuarenta secciones más adelante.

Sin embargo, hay una interpretación alternativa, debida a Hacker y Baker, que ve de un modo bastante diferente las relaciones entre las consideraciones sobre seguir una regla y el argumento contra el lenguaje privado (Hacker y Baker, 1984). Según su interpretación, lo esencial de las consideraciones sobre el seguimiento de reglas es que seguir una regla es una práctica, una costumbre o un modo de comportamiento. Seguir una regla, por ejemplo, la regla para el uso de la palabra «dolor» (la regla involucrada en el concepto de dolor), no es comprobar dentro de uno mismo, por decirlo de algún modo, si la nueva sensación se parece o no a la sensación original y, en caso de que se parezca, pensar que es dolor. Obedecer una regla es asunto de conducta pública, no la operación de un mecanismo privado. Es, por tanto, importante que la sección citada anteriormente comience de hecho: «Por lo que también "obedecer una regla" es una práctica. Y *pensar* que uno está obedeciendo una regla...» Pero si obedecer una regla es una práctica, que debe ser pública sólo en el sentido en que es una

manera de comportarse más bien que de estar controlado por la comunidad, ¿cómo llegar a la conclusión de que el lenguaje privado del solipsista es imposible? ¿Por qué no sería posible una práctica que pudiera ejercerse en soledad, en aislamiento de cualquier comunidad? También podemos preguntarnos cuál sería la respuesta que, de acuerdo con esta interpretación, daría Wittgenstein a las cuestiones de Kripke sobre la objetividad. Ciertamente se trata de cuestiones oportunas que deben ser afrontadas y, al abandonar la interpretación comunitaria, nos privamos de la única posible respuesta.

La primera de estas objeciones a la interpretación Hacker–Baker puede responderse diciendo que, aunque la práctica que constituye conocer y seguir una regla para el dolor podría tener lugar lejos de una comunidad (en una isla desierta) una vez que hubiera sido establecida, el problema no radica en el ejercicio de esa práctica, sino en su misma institución. Es decir, el solipsista no puede conseguir, como pretende, que la práctica sea instituida por el mero hecho de concentrarse en la naturaleza de la sensación original y de inventar un nombre que se refiera a esa sensación y a otras similares a ellas.

Ya hemos considerado algún fundamento para esta conclusión al discutir la interpretación comunitaria. No es sorprendente, dado que ambas interpretaciones comparten el mismo punto de partida: la sensación original no puede servir para fundamentar la práctica (no puede crear el concepto relevante). Pero necesitamos más argumentos, dado que la versión de Kripke dependía estrechamente de su análisis de las relaciones entre pensamientos sobre seguimientos de reglas y el argumento del lenguaje privado. Afortunadamente, Wittgenstein ofrece tales argumentos. El problema no es el encontrarlos, sino el seleccionarlos y ordenarlos de un modo comprensible.

En primer lugar, indica que el proceso de definición ostensiva, por la que el solipsista pretende dar significado al término «dolor» señalando (mentalmente) a una sensación, sólo funciona cuando tenemos previamente un trasfondo de conocimiento conceptual, y no puede usarse para construir tal conocimiento a partir de una página en blanco. Supongamos que señalo a una silla y digo «Por "silla" quiero decir *esto*». Nada de lo que he hecho crea el significado deseado para «silla», a menos que pueda caracterizar además qué es lo que considero relevante en el objeto que acabo de señalar; por ejemplo, podría decir «este tipo de mueble», lo que mejoraría algo las cosas. Pero, en ese caso, necesitaría tener antes el concepto de «mueble». Sucede lo mismo con las sensaciones: si dijera «por "dolor" quiero decir *esto*», el efecto buscado sólo se alcanzaría si tuviera algún modo de distin-

guir las características relevantes de las que no lo son. ¿En virtud de qué han de ser irrelevantes la duración, la intensidad, la localización, la causa y el poseedor del dolor? En general, la definición ostensiva descansa sobre la existencia previa de conocimiento conceptual, y no puede ser la primera fuente de tal conocimiento. Lo que quiere decir que nunca puede existir el tipo de pura definición ostensiva privada que el solipsista y otros necesitan.

En segundo lugar, supongamos que sea necesaria una regla formulable para la palabra «dolor», conteniendo una referencia explícita a la sensación inicial, como en «dolor = sensación como esto». Se supone que la regla es algo que podemos usar para comprobar si la nueva sensación es dolor. Pero, si estamos inclinados a llamarla «dolor», ¿qué contenido tendría el supuesto proceso de comprobar si nuestra inclinación está o no justificada? Wittgenstein dice que lo único que podemos hacer es mirar otra vez para ver si sentimos realmente la inclinación de llamarla «dolor». Pero no hay aquí una comprobación independiente: es más bien como comprar un segundo ejemplar del diario para comprobar si el primero que compramos decía la verdad. La tentación de denominar «dolor» a la nueva sensación es la tendencia de creer que es semejante a la sensación original. No podemos comprobar independiente esta creencia examinando de nuevo la sensación original, no porque no podamos volver a examinarla, sino porque todo lo que podríamos hacer es repetir lo que ya hicimos al pensar que la nueva sensación es relevantemente similar a la antigua.

Lo esencial en este punto no es que las únicas comprobaciones que podemos hacer sobre nuestras inclinaciones son comprobaciones falibles. Si hubiera un proceso real de comprobación no haría falta exigir que fuera infalible; el hecho de que siempre pudiéramos estar equivocados sobre cómo era la sensación original (una especie de escepticismo sobre la memoria) no es, en absoluto, parte del argumento. El argumento no es que la comprobación es falible; si así fuera, siempre podríamos esperar que un número suficiente de recuerdos pudieran apoyarse entre sí, como sugiere Ayer (1954). Lo esencial es, más bien, que la supuesta comprobación independiente es una mera repetición del procedimiento que se supone que comprueba.

Esta conclusión no debe verse afectada por ninguna consideración sobre si el uso de la memoria constituye un nuevo examen de la sensación original o una mera inspección de cómo nos parece ahora que ha sido esa sensación. Cualquiera que sea de esas dos alternativas (no muy claras) la que adoptemos (véase 12.2–3), todavía es cierto que lo que proporciona la memoria no es una comprobación distinta de nues-

tra inclinación a creer que la nueva sensación se parece a la antigua. Todo lo que pueda proporcionarnos la memoria contribuye a nuestro sentimiento de que, al denominar «dolor» esa sensación, continuamos siguiendo la regla del mismo modo. Por lo que no podemos comprobar si seguimos la regla tratando de recordar la sensación original.

Por último, Wittgenstein observa que, si el lenguaje se usa para la comunicación, sus términos no pueden obtener significado a partir de los objetos privados de uno de sus usuarios. Compara ese supuesto escenario con un grupo de personas que tienen cajas; en las cajas hay distintas cosas (algunas pueden estar vacías) y cada persona denomina «escarabajo» al contenido de su caja. Si el término «escarabajo» adquiere significado a partir del contenido de las cajas, no puede usarse para la comunicación con los demás, que, desconociendo la naturaleza de ese escarabajo, no entenderán el término. Si el término tiene un significado independiente de la naturaleza del objeto, puede usarse para la comunicación (quizá signifique «el contenido de la caja»), pero el objeto privado de la caja queda cancelado por irrelevante: una especie de clavija desconectada.

Este último argumento parece de escasa utilidad contra el solipsismo, que no está interesado en la comunicación con otros. Sin embargo, es posible que haya algún modo de adaptarlo para mostrar que el solipsista no puede usar el término «escarabajo» para comunicarse con su yo posterior (por ejemplo, en un diario), dado que lo que da significado al término no puede ser, ni siquiera para él, lo que estaba en la caja (un objeto al que no tiene acceso ahora), sino sólo lo que él crea que había en ella.

Parece, entonces, que el argumento del lenguaje privado tiene una vida independiente con respecto a la interpretación dada por Kripke de las consideraciones sobre el seguimiento de reglas. En vez de decir que no puede haber reglas en ausencia de una comunidad, lo incorrecto en las reglas del solipsista es que no pueden establecerse, tal y como pretende, usando la experiencia privada como una muestra. Y ese extremo es independiente del *número* de usuarios del lenguaje privado. Si, *per impossibile*, el lenguaje pudiera establecerse podría funcionar sin ninguna comunidad que lo controlara. Lo erróneo del lenguaje del solipsista no es que carece de control comunitario; Robinson Crusoe *puede* hablar consigo mismo y llevar un diario en su isla desierta. Pero su lenguaje no es el del solipsista: no se fija por definición ostensiva de sensaciones privadas usadas como nuestras.

¿Qué sucede con la segunda objeción a la interpretación de Hacker y Baker, la de que no pueden responder a los retos de Kripke sobre la

objetividad? Aquí, estos autores adoptan una línea intransigente, negando que Wittgenstein estuviera impresionado en lo más mínimo por tal tipo de cuestiones. Su objetivo principal respecto al escepticismo es el de rechazar sus supuestos, no el de responderlos, y, en este caso, el supuesto es el de que necesitamos alguna base para la objetividad de las prácticas de seguir una regla. Las dudas escépticas sobre la objetividad tienen sentido *dentro* de una práctica; hay respuestas objetivamente correctas a las cuestiones sobre cómo seguir, y, si se nos reta sobre un caso particular, podemos apoyar nuestra elección apelando a la regla («¿Por qué escribiste 20.002?», «Porque me dijiste que continuara añadiendo 2»). Pero las cuestiones sobre la consistencia objetiva de una práctica, realizadas, por así decirlo, desde el exterior de esa práctica, no tienen ningún sentido. (¿Qué me justifica para escribir 20.002 después de 20.000 cuando estamos añadiendo 2? Nada me justifica; eso es lo que denominamos «añadir 2».) Y las cuestiones escépticas que Kripke considera cruciales son todas cuestiones de este tipo. Involucran la separación de la regla respecto a sus instancias y nos preguntan qué nos justifica para considerar que esas instancias lo son de esa regla particular. Pero, si separamos de ese modo las instancias y la regla, no queda regla alguna sobre la que hacernos este tipo de pregunta. Las instancias son la práctica, y la práctica es la regla; una regla está «internamente relacionada» con sus instancias y los intentos del escéptico de establecer una separación entre ellas relevan una confusión sobre qué sea seguir una regla.

Hay en esta actitud de Wittgenstein mucho más que una negativa completa a tomar parte en el juego filosófico. Su posición es que deberíamos pensar en las reglas como prácticas, como formas de actuación, y no suponer que debe existir, para apoyar esas formas de actuar, alguna interpretación interna de la regla que les diga a los que la siguen cómo llevarla a la práctica. Si suponemos tal cosa, ingresaremos en un regreso al infinito; nos encontraríamos a la búsqueda de una nueva interpretación. El seguimiento de una regla en tanto que práctica debe poder ejecutarse, por tanto, sin nada que lo apoye, y debemos intentar alejarnos del sentimiento de vértigo engendrado por el reconocimiento de esa falta de apoyo. De un modo semejante, no debemos permitirnos buscar algo por debajo de la práctica sobre lo que fundamentar el sentimiento de que la practicaba a continuar de un modo objetivamente correcto. La objetividad es parte de la práctica, no algo que la apoye o la fundamente desde afuera. La objetividad es el elemento de la práctica que involucra la comprobación, la reafirmación, el abandono de posiciones previas, las quejas sobre los demás,

etc. Pero nada de todo eso es necesario como fundamento que convierta a esas actividades en algo de algún modo justificado.

5.7. CONCLUSIONES COMUNES

Dadas las diferencias entre las dos interpretaciones del argumento de Wittgenstein, deberíamos poner de relieve sus semejanzas.

En primer lugar, ambas están de acuerdo en que el concepto de «dolor», que nos ha servido de ejemplo a lo largo de toda esta discusión, no es especial en ningún sentido relevante. Sería posible sentir la tentación de decir que el análisis de Wittgenstein podría ser correcto con relación a las sensaciones; pero que no hay cosa tal como una sensación de azul, por lo que no hay ninguna razón por la que un hablante privado no pudiera comenzar dando nombre al modo en que le parecen las cosas antes de pasar al lenguaje mucho más difícil sobre cómo son en realidad las cosas. Las dos interpretaciones comentadas considerarían que ésta es una escapatoria equivocada. No hay ninguna diferencia relevante entre sensaciones y otros posibles «contenidos de conciencia». Podemos comprobarlo volviendo a considerar los argumentos procedentes. Nada de lo que se ha dicho dependía de una característica esencial del dolor. Y, por supuesto, hemos visto que Wittgenstein deseaba mostrar la imposibilidad de que existieran seres que se comportaran de un modo similar al nuestro y que carecieran de sensaciones u otros tipos de experiencias distintas a las nuestras. Y pretendía considerar de un modo similar nuestra creencia inicial de que, aunque usemos el mismo lenguaje y estemos de acuerdo en las palabras con las que describimos los colores de los objetos que nos rodean, todavía podríamos, por todo lo que sabemos, *ver* los objetos de un modo completamente distinto; un objeto que causa en mí lo que denomino una sensación de rojo puede causar en ti lo que yo denominaría una sensación de azul, aunque nunca pudiéramos advertirlo dado que continuaríamos coincidiendo respecto a qué decir y cómo actuar en los distintos medios coloreados (por ejemplo, ambos nos detendríamos ante el semáforo rojo). De acuerdo con Wittgenstein, esta creencia inicial es tan incoherente como la declaración del solipsista de que es capaz de fijar un lenguaje privado, dado que descansa en una enorme separación entre nuestra conducta (incluyendo la conducta lingüística) y la vida mental interior.

En segundo lugar, ambas interpretaciones están de acuerdo en la conclusión de que la pantomima solipsista de concentrarse en una

experiencia y pronunciar una palabra, no importa cuan fervientemente, no puede establecer una regla que el solipsista pueda seguir en un caso nuevo. La naturaleza de la sensación no puede crear una regla, la regla puede crear una naturaleza para la sensación, capacitándonos para concebir distintas maneras en las que esa sensación podría parecerse a otras. Hasta que no tuviéramos semejante concepción, no podríamos describirnos a nosotros mismos la naturaleza de la sensación; sería sólo un *esto* inútil, dado que el concepto de sensación no tiene ningún contenido hasta que no se hayan fijado las reglas.

Esta conclusión es similar a la crítica al fundamentalismo clásico que se hizo en el capítulo 4. Allí se dijo que las creencias del fundamentalista sobre sus propios estados sensoriales, para ser infalibles, deberían tener un contenido mínimo y evanescente; de hecho, equivaldrían sólo a una especie de gesto incomprensible. Esta semejanza entre los dos polos del ataque al fundamentalismo clásico parece reconfortante.

Sin embargo, vale la pena que recordemos que la crítica de Wittgenstein al solipsismo pretendía dirigirse también al programa del empirismo clásico expuesto por el fundamentalismo clásico. Esta observación puede hacerse de dos maneras. En primer lugar, Wittgenstein considera este tipo de fundamentalista como una suerte de escéptico, que admite la dificultad de mostrar que alguna vez estemos justificados para creer que existen otras personas. Pero ve a este fundamentalista como un escéptico tímido que, para ser consistente, debería atreverse a ir más allá; que, de hecho, debería dudar de su comprensión de la proposición de que existen otras mentes, por lo que debería ser un solipsista. Las críticas del solipsismo deben, por tanto, contar como críticas del tímido escepticismo incorporado al fundamentalismo. Si, por el contrario, resistiéramos el intento de Wittgenstein de empujarnos desde el escepticismo tímido al solipsismo, es posible que encontráramos poco convincente el argumento relevante (véase 5.4). Lo que Wittgenstein quiere decir puede expresarse de otro modo. El programa del fundamentalismo clásico de empezar a partir del propio caso y de moverse hacia afuera sufre de todos los defectos radicales que afectan al intento solipsista de fijar un lenguaje en el que describir las experiencias que constituyen su objeto. Ni el fundamentalismo ni el solipsismo pueden realizar el *primer* movimiento. Estas posiciones están minadas por su creencia equivocada de que sabemos lo que significan los términos de experiencia «dolor», «parece rojo» por familiaridad directa con las cosas a las que se refieren.

5.8. PERSPECTIVAS PARA EL FUNDAMENTALISMO

Vimos en 4.3 que hay teorías fundamentalistas distintas a las de la variedad clásica. ¿Qué sucede con ellas si aceptamos los argumentos de Wittgenstein?

El fundamentalismo acepta que hay dos tipos de justificación: la inferencial y la no–inferencial. Este supuesto no se ve afectado por el argumento contra el lenguaje privado. La asimetría involucrada en la distinción entre dos tipos de justificación sólo se convierte en peligrosa si se supone que todas las creencias básicas (justificadas de un modo no inferencial) tratan sobre la naturaleza de los estados sensoriales presentes del creyente. Por lo que hace al argumento del lenguaje privado, el fundamentalismo es todavía posible en la medida en que evite el punto de vista tradicional de que la epistemología es la empresa de comenzar a partir del propio caso y de construir a partir de él.

No será una sorpresa advertir que Wittgenstein es, a su manera, un fundamentalista. Escribe (Wittgenstein, 1969, § 136):

> Cuando Moore dice que sabe tal y tal [por ejemplo, que tiene dos manos (J. D.)], en realidad, está enumerando un conjunto de proposiciones empíricas que afirmamos sin comprobación especial; es decir, proposiciones que tienen un papel lógico peculiar en el sistema de nuestras proposiciones empíricas.

Y esta opinión de que algunas proposiciones (creencias) desempeñan un papel especial en nuestro conjunto de creencias es sintomáticamente fundamentalista.

Wittgenstein no es un fundamentalista ordinario, las creencias que, según su sugerencia, desempeñan ese papel especial para nosotros incluyen creencias como que tengo dos manos, que los hombres no vuelan a la Luna, que el Sol no es un agujero en el cielo, que la Tierra ha existido durante el último siglo y que nuestras manos no desaparecen cuando no les prestamos atención. Está de acuerdo en que podemos preguntar qué justifica esas creencias; pero hacer tal cosa significaría que ya no les concedemos ese *status* especial. El *status* especial que tienen no necesita de fundamentos o de justificación. Se trata tan sólo de que las tratamos como proposiciones que no necesitan de justificación pero que pueden justificar otras. No preguntamos, ni siquiera nos lo planteamos, qué las justifica.

Si, después de los argumentos de este capítulo, Wittgenstein todavía puede ser considerado un fundamentalista, es obvio que aún no está decidido cuáles son las últimas perspectivas para el fundamenta-

lismo. En el capítulo 7 se ofrecerá una razón para rechazar todas las formas de fundamentalismo.

LECTURAS COMPLEMENTARIAS

Wittgenstein (1953, §§ 143-242) sobre seguir una regla y (§§ 243 y siguientes) sobre lenguajes privados.

Hacker (1972, caps. 7-10) proporciona un análisis más detallado de las opiniones de Wittgenstein.

Kripke (1982) es una versión más extensa de Kripke (1981), incluyendo un capítulo sobre Wittgenstein y las otras mentes. Baker y Hacker (1984) atacan explícitamente la interpretación de Kripke.

El conductismo extremo se asocia normalmente a B. F. Skinner. Una de las discusiones menos desfavorables al conductismo de Skinner puede encontrarse en Dennett (1978, cap. 5).

Ryle (1949) representa otro intento de encontrar un camino entre el conductismo extremo y la separación mental/físico.

Desafortunadamente, la falta de espacio ha impedido cualquier comentario sobre las opiniones positivas de Wittgenstein en epistemología, para las cuales véase Wittgenstein (1969a). La contribución de Wittgenstein en este terreno podría haber tenido una resonancia mucho mayor en el presente capítulo.

La noción de criterio es discutida por Albritton (1959). Un punto de vista novedoso respecto al papel de los criterios se encuentra en McDowell (1982), al que responde Wright (1984).

6. TEORÍAS EMPIRISTAS DEL SIGNIFICADO

6.1. LA RELEVANCIA DE LAS TEORÍAS DEL SIGNIFICADO PARA LA EPISTEMOLOGÍA

Existe una íntima conexión entre la epistemología y la teoría del significado que ya se ha vislumbrado en los capítulos anteriores. En el capítulo 1, por ejemplo, considerábamos el efecto que podrían tener distintas teorías de la comprensión sobre ciertos argumentos escépticos; y una teoría de la comprensión es sólo una teoría del significado con otro nombre. Del mismo modo, la respuesta antirrealista al escéptico descansa en el punto de vista de que no podemos comprender proposiciones cuyo significado es tal que expresan (o pretenden expresar) hechos que trascienden toda evidencia. Vimos también, en 1.2 y en 5.4, que los argumentos escépticos más poderosos atacan tanto la comprensión como el conocimiento. Por último, en el capítulo 5 se utilizó una teoría del significado como forma de desacreditar cierto programa epistemológico. Ha llegado el momento de detenernos para ver exactamente cómo ha sido posible todo esto.

La crítica de Wittgenstein al solipsismo, o al fundamentalismo más consistente, era que el solipsista no puede, a pesar de sus pretensiones, desarrollar ningún lenguaje. La crítica depende de una consideración sobre qué es el ser competente en un concepto, sobre qué es conocer el significado de una palabra, sobre qué es conocer las reglas para la aplicación de esa palabra: las reglas que se supone que nos llevan de una instancia a otra. Wittgenstein argumentó que el solipsista no puede elaborar esas reglas del modo pertinente y, por tanto, que no puede construir un lenguaje. De ello concluíamos que cierto programa epistemológico nunca podría tener éxito. De modo que podría decirse que el error básico que subyace al fundamentalismo clásico es un error en la teoría del significado.

Sabemos que existen otras formas de fundamentalismo distintas a la clásica. En este punto, ¿podemos aislar una teoría del significado que, si fuera correcta, nos forzara a adoptar alguna forma de fundamentalismo? Si tal cosa fuese posible, nos sería útil de dos maneras.

En primer lugar, podríamos utilizarla para persuadirnos de que los errores que Wittgenstein encontró en el fundamentalismo clásico no contaminan realmente todas las formas de fundamentalismo. En segundo lugar, podríamos discernir lo correcto y lo erróneo en tal teoría del significado, y usar estas conclusiones para determinar nuestra actitud hacia el fundamentalismo en general, de modo que no necesitaríamos considerar, una a una, todas las posibles variedades del mismo.

En último término, ésta será mi estrategia. Argumentaré que los atractivos del fundamentalismo, sean los que sean, no derivan de la teoría del significado que subyace al mismo; de hecho, la forma más aceptable de teoría del significado carece de los rasgos característicos del fundamentalismo y, en realidad, sirve de apoyo a una epistemología alternativa, el coherentismo.

6.2. EL EMPIRISMO LÓGICO Y LA EVIDENCIA DE LOS PROPIOS SENTIDOS

En 4.3 sugerí que el fundamentalismo debe definirse por su respuesta al argumento del regreso de justificaciones:

F^1: Hay dos formas de justificación, la inferencial y la no–inferencial.

Pero hay otra tendencia característica del fundamentalismo, el empirismo. El fundamentalismo, desde un punto de vista general (aunque aquí debamos excluir a Wittgenstein), puede ser considerado como la expresión del punto de vista empirista de que la verificación y la justificación, que discriminan si algo es o no verdadero y respaldan las declaraciones propias sobre lo que es verdadero, deben descansar, en último término, en la evidencia de los propios sentidos; puede que no en primera instancia, pero sí al final del trayecto. ¿A qué más podríamos apelar, para que nos diga si algo es verdad, que a la evidencia de nuestros sentidos? He aquí un reto empirista que parece sensato. De acuerdo con él, es la evidencia de nuestros sentidos el punto de partida siempre que necesitamos proveernos de justificaciones para nuestras creencias.

Llegados aquí, es esencial percibir que la evidencia de los propios sentidos no es sólo aquello a lo que apelamos cuando tratamos de justificar y verificar. Es también el punto de partida para el aprendizaje

del lenguaje. También en este respecto podríamos preguntarnos qué otro punto básico de apoyo podríamos encontrar que no fuera lo que es dado, es decir, la evidencia de nuestros sentidos. De modo que los empiristas consideran que esta noción de evidencia de los propios sentidos es la básica, no sólo en epistemología, sino también en teoría del significado. Una teoría empirista del significado será una que consagre esta dependencia de todo aprendizaje del lenguaje, y, con ello, de todo lenguaje significativo, sobre la evidencia de los propios sentidos. Llevados por la idea de que la evidencia de los propios sentidos es básica tanto en epistemología como en teoría del significado, positivistas como Ayer propusieron, como una teoría del significado, lo que denominaron el principio de verificación del significado empírico [véanse Ayer (1946) y Schlick (1936)]:

PV: un enunciado tiene significado empírico si y sólo si su verdad establecería alguna diferencia respecto a la evidencia de nuestros sentidos,

que (dado que este punto de vista considera que un enunciado es verificable si y sólo si su verdad establece alguna diferencia respecto a la evidencia de nuestros sentidos) es equivalente a:

PV[1]: un enunciado tiene significado empírico si y sólo si es verificable.

Y a partir de este análisis del significado empírico surge fácilmente un análisis de en qué consiste el que alguien comprenda un enunciado o conozca su significado:

PC: *a* conoce el significado de *p* si y sólo si *a* sabe cómo verificar *p*.

Lo que podría ampliarse del siguiente modo:

PC[1]: *a* conoce el significado de *p* si y sólo si *a* sabe qué diferencia establecería la verdad de *p* respecto a la evidencia sensorial de *a*.

De donde derivamos fácilmente un análisis del significado de cualquier enunciado particular E:

PS: el significado de E es la diferencia que la verdad de E establecería respecto a las propias evidencias sensoriales.

Lo que se interpretó a veces, y de un modo más bien desafortunado, como:

PS[1]: el significado de E es su método de verificación.

Como puede verse, estos tres principios son aceptablemente empiristas. De hecho, PV era y es el núcleo de la posición conocida como empirismo lógico o positivismo lógico.

¿Qué quiere decir la palabra «verificar» en esos principios? Tradicionalmente, se traza una distinción tradicional entre verificación fuerte y verificación débil [véase Ayer (1946), p. 9]. La verificación fuerte es verificación concluyente; un enunciado es verificable de un modo concluyente si, una vez que tengamos la mejor evidencia posible para él, no queda ninguna posibilidad de que sea falso. La verificación débil no llega a ser concluyente. Un enunciado que es verificable de un modo débil puede ser confirmado o desautorizado por otros enunciados que sean verificables de un modo concluyente, sin que él mismo sea verificable de ese modo; es decir, enunciados que son verificables en sentido fuerte pueden contar como evidencia a favor o en contra suya.

Esta distinción puede ayudarnos a dar la mejor versión del empirismo lógico. Podríamos insistir en que «verificable» en VP[1] significa «verificable de modo fuerte». Pero tal cosa eliminaría como no significativos tantos de nuestros enunciados que sería autorrefutante. Por otra parte, no es normal que los empiristas sientan la inclinación de suponer que no hay, en absoluto, enunciados verificables en sentido fuerte. Más bien, consideran que los enunciados que sólo informan sobre la propia evidencia sensorial, sea la que sea, pueden ser verificados concluyentemente cuando los propios sentidos proporcionan de hecho la evidencia en cuestión. Parece, pues, que la forma más plausible de empirismo lógico sostiene que los enunciados pueden clasificarse en dos tipos, los que son verificables de un modo fuerte y los que no tienen ellos mismos ese tipo de verificabilidad, aunque puedan ser confirmados o desautorizados por otros enunciados que sí son verificables en sentido fuerte.

Podemos ver ahora lo estrecha que es la relación entre empirismo lógico y fundamentalismo. La pretensión del fundamentalismo de que hay dos tipos de justificación, la inferencial y la no inferencial, se refleja en el supuesto del empirismo lógico de que todos los enunciados significativos son verificables en sentido fuerte o en sentido débil.

¿Qué enunciados son verificables de un modo concluyente, o, en

otras palabras, qué enunciados reflejan la evidencia de los propios sentidos? Por usar la terminología de Ayer, ¿qué enunciados son «enunciados observacionales» [*observation statements*], y cuáles informan sobre estados de cosas que están más allá de lo que, en sentido estricto, puede observarse? Los empiristas lógicos discrepan respecto a esta cuestión, del mismo modo en que discrepan los fundamentalistas respecto a la naturaleza de las creencias básicas. Ayer, por ejemplo, adopta la posición clásica de que los enunciados de observación son los que describen la naturaleza de los propios estados sensoriales. Pero no es ésta una posición obligatoria; lo que nos viene bien, dado que considero que los argumentos de los capítulos 4 y 5 ya la han refutado. Quine escoge la alternativa de decir que la evidencia sensorial propia versa no sobre lo interno al observador, sino sobre lo que es externo a él, es decir, la presencia de cierto estímulo (público) [Quine (1975a), p. 73]. Un enunciado observacional se hace como respuesta a ciertos estímulos y es verificable en sentido fuerte por apelación a tales estímulos. De modo que la noción de un enunciado de observación, que no informa de nada más que de la evidencia de los sentidos, es una noción controvertida entre los empiristas lógicos. Lo que no se discute es la distinción entre verificación fuerte y débil y su importancia para la teoría del significado. La versión de Quine de empirismo lógico daría apoyo a una variante de fundamentalismo que no sería vulnerable a los argumentos del capítulo 5, dado que, para él, los enunciados observacionales básicos para la justificación, la verificación y el aprendizaje del lenguaje no versan sobre la naturaleza de los estados sensoriales propios.

6.3. TRES TEORÍAS VERIFICACIONISTAS

¿Cuál es exactamente la relación entre enunciados de observación y los otros, que forman la enorme mayoría? Hemos sugerido que los últimos pueden ser confirmados o desautorizados por los primeros; pero ¿cómo han de estar relacionados para que tenga lugar tal confirmación? Debemos dar un análisis al respecto que sea consistente con el principio de verificación. De hecho, esto quiere decir que nuestra respuesta equivaldría a un análisis del carácter significativo de los enunciados no–observacionales, y de lo que hace que uno de esos enunciados tenga un significado distinto al de otro de ellos. Consideraremos tres respuestas a la cuestión, que forman una especie de espectro.

FENOMENALISMO

El fenomenalismo, como teoría del significado de los enunciados no–observacionales, sostuvo, en un principio, que tales enunciados son equivalentes en significado a una lista (probablemente muy larga) de enunciados (unidos por una conjunción) sobre lo que se observaría bajo circunstancias diferentes. Por ejemplo, lo que tiene que suceder para que haya una rosa roja en esta habitación oscura es que se dé el caso de que, si encendiera la luz, haría una determinada observación, de que, si me desplazara hacia un determinado lugar, haría otra observación diferente, y de que, si tú entraras en la habitación, observarías tales y tales cosas, etc. Esta teoría tiene un enorme atractivo tanto en metafísica como en epistemología, dado que fenomenalistas como Ayer tendieron a considerar que los enunciados observacionales informan sobre la naturaleza de los estados sensoriales del observador. Si esto fuera así, nos encontraríamos con el bonito resultado de que, en vez de haber dos tipos de cosas radicalmente diferentes, estados sensoriales y objetos materiales, sólo hay un tipo, los estados sensoriales, y todos los otros tipos que se pudieran suponer no serían más que complejos de cosas reales y posibles del primer tipo. Ésta es la ventaja metafísica del fenomenalismo. El atractivo epistemológico, por supuesto, radica en que es genuinamente posible el conocimiento del mundo «externo» de objetos materiales. Tales objetos no yacen más allá de nuestras capacidades, como podría sugerir el escéptico; dado que podríamos esperar que cada miembro del conjunto de enunciados sobre lo que se observaría fuera verificado de un modo concluyente, y, en tal caso, no quedaría ninguna posibilidad de que el enunciado sobre objetos materiales fuera falso. Así pues, el fenomenalismo de este tipo sería una forma de antirrealismo (véase 1.4).

Si cada miembro del conjunto de enunciados observacionales que, en conjunto, determinan el significado de un enunciado no–observacional, pueda especificarse y verificarse de un modo preciso, ese enunciado no–observacional tendría, de acuerdo con el principio de verificación, su propio significado determinado y, en ciertas circunstancias, estarían determinadas su verdad o su falsedad. Esta situación constituye el ideal del fenomenalista: todos los enunciados no–observacionales serían verificables en sentido fuerte y las dudas escépticas sobre ellos serían imposibles. Pero las cosas no son así, por dos motivos. En primer lugar, parece posible que, en la mayoría de los casos, no puedan verificarse todos los enunciados condicionales sobre lo que se observaría en ciertas circunstancias, porque habría casos en los que

la verificación de uno eliminara cualquier posibilidad de verificar otro. Si sólo pueden verificarse algunos de ellos, el enunciado no–observacional no será más que verificable en sentido débil. En segundo lugar, no está claro que los enunciados condicionales sean verificados en sentido fuerte por el mero hecho de que se muestre la verdad conjunta del antecedente y del consecuente. El lector sabe leer castellano y está leyendo este libro, pero no es cierto que, si sabe leer castellano, entonces está leyendo este libro. De modo que, si el significado de un enunciado no–observacional está dado por un conjunto de enunciados sobre lo que se observaría si..., no es probable que sean verificables en sentido fuerte. (Aunque pudieran ser refutables en sentido fuerte.)

Esto quiere decir, por supuesto, que las ventajas contra el escéptico no son tan espectaculares como se pretendía. Sin embargo, algo sí ganaríamos, dado que no pretenderíamos conocer, basándonos en la experiencia, la naturaleza de un mundo independiente de la experiencia. El fenomenalismo reduce la separación entre el mundo de experiencia y nuestra experiencia de él.

Hasta ahora se ha hablado del fenomenalismo como si todos los fenomenalistas estuvieran de acuerdo en que los enunciados de observación informan sobre la naturaleza de los estados sensoriales del propio informador. Pero tal fenomenalismo sólo es una de las formas de reduccionismo. En vez de ello, podríamos sostener que los enunciados de observación informan sobre la naturaleza del mundo cercano, y definir todos los otros enunciados en términos de éstos (considerándolos como condicionales sobre la localización del observador). Esto constituiría una teoría también reduccionista, compatible con las exigencias del empirismo y del principio de verificación, pero quizá careciera de algunos de los atractivos epistemológicos y metafísicos de la teoría tradicional.

Un problema crucial para el fenomenalismo de cualquier tipo es el hecho de que pocos filósofos han intentado (y ninguno lo ha conseguido ni siquiera remotamente) mostrar los detalles del tipo de reducción que nos propone. Tan pronto como se intentó detallarla, la teoría se demostró impracticable. Independientemente de otras dificultades, parecía implausible suponer que un enunciado sobre la rosa roja en la oscuridad hubiera de tener unas consecuencias observacionales completamente precisas. De modo que, además de ser probablemente de una longitud infinita, la lista de enunciados condicionales de observación habría de tener un contenido más bien vago. Como dijo Ayer, «lo que se requiere para verificar un enunciado sobre un objeto material

no es tanto la ocurrencia de uno u otro contenido sensorial bien precisos, como la mera ocurrencia de uno u otro de los contenidos sensoriales abarcados por un rango más bien indefinido» [Ayer (1946), p. 12]. Y con esta admisión el fenomenalismo pierde una de sus ventajas. Dado que, en la medida en que el rango implicado es impreciso, en esa misma medida los enunciados no–observacionales tienen un significado indeterminado, y el rango de circunstancias en las que su verdad o su falsedad están determinadas se ve enormemente reducido, quizá a la pura nada. Lo que esto significaba era el abandono de la teoría reductiva original, al no poder pretenderse que un enunciado no–observacional fuera exactamente equivalente en significado a un conjunto de enunciados de observación, por más complejo y condicional que fuera el conjunto en cuestión.

LA REVISIÓN DE CARNAP

Rudolf Carnap fue quizá el único filósofo que intentó una reducción completa de los enunciados no–observacionales a complejos explícitamente elaborados de enunciados de observación. Sus intentos fracasaron, y llegó a la conclusión de que lo mejor que se podía hacer era especificar, en la medida de lo posible, qué enunciados observacionales estaban implicados por ciertos enunciados no observacionales [véase el prefacio a Carnap (1967) y Quine (1969), p. 77]. Esto nunca produciría nada semejante a una reducción de unos a otros, pero Carnap supone que todavía nos está permitido declarar que el concepto de objeto material puede reducirse a «conceptos autopsicológicos», aquellos relativos a la naturaleza de los propios estados sensoriales.

Pero si todo lo que podemos hacer, cuando tratamos de especificar los significados de los enunciados no–observacionales, es formular algunas implicaciones observacionales de esos enunciados, parece que siempre habrá algunos aspectos de su significado que se nos escaparán y quedarán sin especificar. Y esto sólo incrementa el grado de indeterminación de los enunciados no–observacionales. Da un sentido legítimo a la idea de que un enunciado verificable en sentido débil sólo puede ser confirmado o desautorizado, sin llegar nunca a ser verificado concluyentemente. Por otra parte, admitimos que, para decir que tal enunciado tiene su verdad determinada, debemos aceptar la existencia de hechos en virtud de los cuales es verdadero, y que tales hechos están más allá de cualquier posibilidad de verificación. El verificacionista que, por ser un antirrealista coherente, se niegue a admitir

la posibilidad de este tipo de hechos deberá decir que un enunciado así nunca tendrá determinada su verdad, aunque pudiera estar determinada su falsedad si se observara que una de sus consecuencias es falsa.

QUINE

La posición de Quine es más radical. Argumenta que ni siquiera podemos esperar la especificación de las oraciones observacionales que son consecuencia de unas oraciones no–observacionales determinadas.

La teoría de Quine parte en este punto de tres supuestos bien distintos. El primero es su tesis de la indeterminación de las teorías por la evidencia (los datos). Independientemente de las evidencias que podamos tener, siempre habrá teorías diferentes que explicarán y asimilarán los datos de un modo igualmente satisfactorio. Por ejemplo, es compatible con los resultados de todas las mediciones reales y posibles mantener tanto que el universo tiene un tamaño constante como que se está expandiendo a una velocidad constante. No hay ninguna teoría que esté implicada por sus datos. Distintas teorías pueden tener las mismas consecuencias observacionales.

El segundo es su pretensión de que las oraciones no–observacionales se enfrentan con el tribunal de la experiencia no individualmente, sino en grupos. Quine afirma que esta tesis, que se debe a Pierre Duhem, la podemos denominar la tesis de Duhem [Quine (1969), p. 80]. Duhem sugiere que las oraciones individuales que no son de observación no pueden ser verificadas ni refutadas de un modo concluyente por medio de la observación, por la evidencia de nuestros sentidos. La razón es que tales oraciones no se dan aisladas, en una especie de limbo; se dan como parte de una teoría más general. Por ello, debemos decidir dónde alterar la teoría cuando las cosas no funcionan en el nivel observacional. Siempre habrá más de una manera en que pueda hacerse esta rectificación: las experiencias recalcitrantes nunca nos podrán forzar a alterar precisamente una oración no–observacional determinada. Del mismo modo, las oraciones no–observacionales individuales nunca se ven confirmadas por la experiencia. La experiencia puede confirmar teorías, y, por ello, confirmar también las oraciones con las que se edifican esas teorías, pero no puede confirmar esas oraciones individualmente y de un modo directo. Independientemente de lo bien que vaya la experiencia, sólo puede confirmar una oración no–observacional a la luz de la teoría que la

rodea; cambiemos la teoría, y la oración no–observacional no podrá ser confirmada en modo alguno por la experiencia.

Estos dos supuestos son distintos. El segundo afirma que las oraciones individuales que no son de observación no están sujetas, por sí mismas, a verificación o refutación por parte de la experiencia. El primero afirma que las cosas que pueden verificarse, las teorías, nunca pueden ser verificadas de un modo concluyente. (Por supuesto, sí pueden ser refutadas de un modo concluyente.)

El tercer supuesto quineano ya ha sido considerado. Es la teoría empirista del significado PS:

PS: el significado de E es la diferencia que establecería la verdad de E respecto a la propia evidencia sensorial.

Quine dice de esta teoría que «no hay otra opción más que la de ser empiristas respecto a la propia teoría del significado lingüístico» [Quine (1969), p. 81].

Quine utiliza estos tres supuestos para argumentar en contra de la idea de que las oraciones individuales no–observacionales tengan por sí mismas un significado separado. Por PS, al especificar las consecuencias observacionales de E, lo que especificamos es el significado de E. Pero, por la tesis de Duhem, ninguna oración no–observacional individual tiene sus propias consecuencias observacionales. Si las tuviera, cuando se diera el caso de que esas consecuencias se vieran contradichas, sabríamos exactamente qué deberíamos rectificar. Pero siempre tenemos un margen de elección a este respecto, por lo que ninguna oración tiene consecuencias observacionales por sí misma. Lo que quiere decir, por PS, que ninguna oración no–observacional tiene por sí misma un significado determinado; dado que no hay nada que sea *la* diferencia que *su* verdad establecería respecto a la observación.

El significado de una oración dada, en la medida en que tenga uno, no es una determinada característica que lleve incorporada a ella. Dado que el significado es asunto de consecuencias observacionales, y tales consecuencias pertenecen a las teorías y no a las oraciones, la conclusión de Quine se convierte en inevitable: el significado pertenece a las teorías más que a las oraciones. Y, dado que, como sucede en el caso de las oraciones, una teoría parcial siempre puede contraponerse a otra, debemos acabar concluyendo, con Quine, que «la unidad de significado empírico es la totalidad de la ciencia».

Esta conclusión puede volverse a expresar de dos maneras distintas, ambas necesarias para discusiones ulteriores. Quine sostiene, en

contra de Carnap, que una oración no–observacional no tiene sus propias consecuencias observacionales [Quine (1953), pp. 40-41]. Lo que demuestra que no hay nada como el significado de la oración, considerada en sí misma. En la medida en que podemos decir qué significa, nuestra respuesta dependerá de la naturaleza del resto de la teoría que rodea a la oración. Así pues, no hay ningún objeto determinado al que podamos denominar el significado de esa oración.

De modo que, en el nivel no–observacional, el significado de las oraciones está indeterminado. Ésta es la tesis de *la indeterminación del significado oracional*.

Respecto a las oraciones no–observacionales, la teoría del significado de Quine puede denominarse *holista*. El atomismo, opuesto al holismo, sostiene que cada oración tiene su propio significado, que queda incorporado a ella a través del cambio de teorías. El holismo es el punto de vista de que los significados de las oraciones son interdependientes, de modo que lo que significa una depende de los significados de las otras, y puede verse alterado por un cambio en otras zonas. El portador del significado es primariamente la totalidad de una teoría, no sus partes, dado que es la totalidad de la teoría la única que tiene consecuencias observacionales. Sólo la totalidad de la teoría puede verse refutada concluyentemente por la experiencia desfavorable.

Quine utiliza la posición que acabamos de alcanzar para argumentar, además, que ninguna oración en nuestra teoría es completamente inmune a la revisión [Quine (1953), pp. 41-43]. Podemos considerar toda una teoría como si fuese una esfera, con oraciones observacionales en la periferia y oraciones no–observacionales en el interior. Comenzando con comparaciones relativamente mundanas sobre los objetos materiales, desplazándonos a través de oraciones más teóricas hasta las leyes de la física, hasta llegar a las leyes de la lógica, en el centro. Si hubiera alguna anomalía en la periferia, sabríamos qué oraciones observacionales deberíamos revisar, y deberíamos hacer también revisiones en el interior, dado que la teoría se habría demostrado falsa. La respuesta normal es la de buscar algún arreglo relativamente menor cerca de la periferia: si no encontramos el pastel donde esperamos encontrarlo, tendemos a suponer que alguien se lo ha comido, y no pensamos que se ha evaporado. Lo hacemos así porque este reajuste no afecta a la estructura general de nuestra teoría. Las oraciones cercanas al centro se resisten a la revisión más que las oraciones de la periferia. Las leyes de la lógica, situadas en el mismo corazón de la teoría, son las más resistentes de todas. Necesitaríamos de circunstancias extraordinarias antes de abandonar leyes como la de no–contradic-

ción. Pero no hay ninguna oración completamente inmune a la revisión. Incluso las leyes de la lógica pueden *revisarse* como respuesta a la investigación empírica. Es posible que el cambio más simple en una teoría, dadas ciertas observaciones extrañas, fuera un cambio estructural en su mismo centro más bien que alteraciones exteriores. Por ejemplo, se ha sostenido que el principio de indeterminación de Heisenberg es una razón para rechazar la ley de tercio excluso.

Esta conclusión es el último remache del ataúd de la noción de significado oracional. Dada esa noción, podría mantenerse que la mayoría de las oraciones, si verdaderas, se convierten en verdaderas por una combinación de su significado y del comportamiento del mundo. Necesitamos conocer algo más que su significado para saber si son o no verdaderas. Tales oraciones se denominan tradicionalmente «sintéticas». Pero también puede haber, y tradicionalmente se ha sostenido que es así, oraciones que sean verdaderas sólo en virtud de su significado. Y tales oraciones analíticas nunca serían revisables. A menos que cambie su significado, no hay ninguna posibilidad de que se conviertan en falsas. La posición de Quine tiene la consecuencia de que no existen tales oraciones analíticas. El sentido en el que podemos hablar del significado de una oración individual no está lo suficientemente determinado como para que sea posible la verdad no-revisable de ciertas oraciones en función de ese significado. No existen las verdades analíticas.

Así pues, el argumento general de Quine es que la teoría empirista del significado PS, conjuntamente con la tesis de Duhem y la de la indeterminación de una teoría por los datos, echan por tierra la noción de significado oracional que se suponía que PS explicaba. Pero esto no es un argumento en contra de PS. Los empiristas deben aceptar PS. La situación la salva la afirmación holista de que la unidad de significado empírico es la totalidad de la ciencia.

LECTURAS COMPLEMENTARIAS

El ensayo clásico de Quine «Epistemology naturalised» [«Epistemología naturalizada»] se encuentra en Quine (1969, pp. 69-90).

Las dudas de Quine sobre la analiticidad fueron expresadas por vez primera en su «Two dogmas of empiricism» [«Dos dogmas del empirismo»], ensayo 2 en Quine (1953). Los argumentos de Quine los discute Putnam (1975, ensayo 2).

La introducción de Ayer (1946) a la segunda edición y el capítulo 1 constituyeron una formulación, poco matizada pero influyente, del positivismo lógico como expresión de empirismo extremo.

Quine (1975a) constituye «una sumaria formulación de mi actitud hacia nuestro conocimiento de la naturaleza».

Føllesdal (1975) constituye un valioso análisis de la epistemología de Quine y su relación con su teoría del significado, por un «aficionado» *.

Berlin (1939) es un buen ejemplo del tipo de respuesta que generó Ayer (1946).

Ayer (1969, pp. 240-243) es una expresión más completa de fenomenalismo. C. I. Lewis (1946, cap. 8) argumenta también en favor del fenomenalismo. Chisholm (1948) critica el fenomenalismo como teoría del significado.

* En castellano en el original. *(N. del T.)*

7. HOLISMO E INDETERMINACIÓN

7.1. LA INDETERMINACIÓN DE LA TRADUCCIÓN

De acuerdo con Quine, el significado oracional está indeterminado. No hay ningún hecho que determine lo que significa una oración individual. Una consecuencia de ello parece ser la de que, del mismo modo que el significado de una oración individual está indeterminado, también está indeterminada la idea de que dos oraciones tengan el mismo significado. La traducción es la empresa de buscar, para una oración dada, otra oración con el mismo significado. Por tanto, la traducción está indeterminada.

Esto no quiere decir sólo que la corrección de la traducción no está completamente determinada por todos los datos posibles. Por supuesto, nuestros datos son siempre incompletos, nunca determinan una traducción como concluyentemente correcta en contraposición a todas sus alternativas que serían así concluyentemente erróneas. Lo que querría decir que la corrección de la traducción siempre trascendería toda evidencia, mientras que sí habría hechos determinando la corrección de una y la incorrección de otras, sin que nunca pudiéramos estar en situación de distinguirlos.

No es ésta la conclusión de Quine. Después de todo, ¿qué podría haber que determinara que una de las traducciones fuera correcta y las otras erróneas? La única respuesta posible está en el significado de la oración original. Pero tal significado no está lo bastante determinado como para fijar, entre dos traducciones alternativas, la corrección exclusiva de una de ellas, aunque nunca siquiera pudiéramos decidir cuál.

Lo que es más, esto no quiere decir que la traducción sea imposible. Podríamos sostener que el significado de una oración dada es lo suficientemente rico e idiosincrásico como para no poder ser capturado nunca por otra oración de un lenguaje diferente. Según este punto de vista, ninguna traducción pasaría de ser una mera aproximación. Pero lo que Quine quiere decir es que no hay nada determinado en el significado de la oración original a lo que deban aproximarse las traducciones. En vez de insistir en la posibilidad de que ninguna traduc-

ción sea adecuada, deberíamos reconocer lo que significa en realidad la indeterminación; las distintas traducciones que son igualmente buenas son lo mejor a lo que podría aspirar cualquier posible traducción. Representan lo que es una traducción correcta. De modo que, en vez de concluir que no hay traducciones correctas, debemos aceptar que siempre habrá más de una. Dos oraciones diferentes en un lenguaje pueden ser ambas traducciones correctas de una oración única en otro.

Podríamos tratar de rechazar esta conclusión pensando que, para muchas oraciones del inglés, parece existir una traducción al castellano cuya corrección está completamente determinada. Quine lo aceptaría, pero mantendría que la razón, en este caso particular, es la existencia de un esquema general no discutido para la traducción inglés–castellano. Con tal esquema, y dentro de él, existe una traducción determinada. Pero lo indeterminado es la elección del esquema. Siempre habrá más de un candidato igualmente adecuado.

Quine no parte de este tipo de caso familiar, sino de lo que denomina traducción *radical* [Quine (1960), cap. 2], Se trata de la tarea, supuestamente ejecutada por un lingüista de campo, de escribir un «manual de traducción» para un lenguaje completamente desconocido, sin ninguna ayuda de hablantes bilingües, diccionarios previamente elaborados o esquemas generales compartidos. (Un manual de traducción es una especie de diccionario cuyas entradas son oraciones en vez de palabras.) El traductor radical se enfrenta con la tarea de escribir ese manual sobre las evidencias que le son accesibles; en pocas palabras, sobre la evidencia de las oraciones que emiten o a las que asienten los «nativos» y los contextos de esas emisiones y esos asentimientos. Intenta establecer una relación entre ambas cosas, tratando de responder a la cuestión crucial: ¿en qué circunstancias observables asentirán los nativos a «p»? La traducción debe estar indeterminada porque ninguna cantidad de evidencia disponible garantizaría nunca que la traducción ofrecida fuera la única correcta. En último término, podremos alcanzar la seguridad de que una traducción es lo bastante correcta, pero nunca la certeza de que no existe otra distinta e igualmente correcta.

El argumento hasta este punto ha sido el de que, dado que el significado está indeterminado al nivel no–observacional, la traducción también debe estarlo. Encontraremos motivos para sentirnos incómodos con él. En 7.4 se ofrecerán argumentos que intentan establecer directamente la indeterminación de la traducción. La indeterminación del significado se seguirá entonces de la indeterminación de la traducción, más que al contrario. Mientras tanto, consideraremos un área en

la que Quine no desea aseverar la indeterminación, el área de las oraciones observacionales.

7.2. QUINE COMO FUNDAMENTALISTA

Ya hemos visto que el verificacionismo, la adopción del PV[1], es una expresión natural de fundamentalismo. Vimos que la distinción del fundamentalismo entre las creencias justificadas inferencialmente y las justificadas no–inferencialmente se reflejaba en la distinción verificacionista entre enunciados verificables de un modo fuerte y los verificables de un modo débil. El fundamentalista, en este punto, insiste en una *asimetría*; está en posesión de una teoría bipolar, sosteniendo que hay dos tipos de justificación. Las creencias no–básicas están justificadas en términos de las básicas, y las básicas están justificadas de algún otro modo. La asimetría radica en el hecho de que la justificación funciona de una única manera, de lo no básico a lo básico. El verificacionismo reduplica esta asimetría en la teoría del significado.

Quine nos ofrece una teoría rica en asimetrías del mismo tipo. A pesar de su posición, holista en términos generales, en cuestiones de significado y de justificación al nivel no–observacional, insiste en el tipo de asimetría entre lo observacional y lo no–observacional que caracteriza al fundamentalismo. Estas asimetrías giran, como sería de esperar, en torno a la noción de oración observacional. La noción, de acuerdo con Quine (1969, pp. 88-89), es

> fundamental en dos conexiones [...]. Su relación [...] con nuestro conocimiento de lo que es verdadero es, con mucho, la tradicional: las oraciones observacionales son el almacén de evidencia para las hipótesis científicas. Su relación con el significado es también fundamental, dado que las oraciones observacionales son las que estamos en posición de aprender a comprender en primer lugar, tanto como niños como en calidad de lingüistas de campo [...]. Ofrecen la única puerta de entrada al lenguaje.

La epistemología de Quine es, en este punto, una consecuencia de su teoría del significado y la traducción. El lingüista de campo que emprende la traducción radical se pregunta bajo qué condiciones los nativos asentirían a una determinada oración ocasional O, y bajo qué condiciones disentirían. Con la mayoría de las oraciones, en particular las oraciones no–observacionales, no aparecerá ninguna respuesta clara porque los nativos no se comportarán de la misma manera en las mismas condiciones. Podríamos explicar esto diciendo que el que los nativos asientan o no a O dependerá normalmente no sólo de la situa-

ción a la que se encaran, sino también de las otras creencias con las que se enfrentan a la situación. Pero esto sólo sería un caso particular del hecho general de que, con la mayoría de las oraciones, es posible que dos personas las comprendan sin coincidir sobre su verdad. Sin embargo, hay otras oraciones sobre las que no surge este problema. Son aquellas tales que cualquiera que las entienda estará de acuerdo en su valor de verdad independientemente de las circunstancias. En estos casos especiales, podemos confiar en poder especificar completamente el significado en términos de condiciones de asentimiento y disentimiento. Lo que quiere decir que, para tales oraciones, el significado está determinado y la traducción también lo está. Podemos esperar que encontraremos una oración en otro lenguaje con condiciones de asentimiento exactamente coincidentes. Y, si está determinada la traducción, el significado debe ser atomista más bien que bolista. Cada oración debe tener su propio significado, si ese significado puede reflejarse exactamente en una oración de otro lenguaje. Las oraciones de este tipo son oraciones observacionales.

No se trata tan sólo de que, de hecho, algunas oraciones sean de este tipo especial. Si eso fuera todo, se nos podría perdonar que nos dejara insatisfechos; podríamos pensar que no se nos han dado demasiadas razones para suponer que existan oraciones tales que todos los que las entiendan estarán, sean cuales sean las circunstancias, de acuerdo sobre su valor de verdad. Recordemos que, para Quine, las oraciones observacionales no informan de sucesos privados como las sensaciones. Informan de que se dan ciertos estímulos sensoriales, y estos estímulos se consideran públicamente accesibles: el mismo estímulo puede suceder a más de una persona. (Decir que dos oraciones pueden tener condiciones coincidentes de asentimiento y disentimiento es lo mismo que decir que pueden ser estimulativamente sinónimas.) ¿Existen, de hecho, tales oraciones? Quine ofrece el ejemplo de oraciones de una sola palabra, como «Rojo» y «Conejo», o «La marea ha bajado», aunque afirma que, en sentido completamente estricto, sólo sería admisible la primera de éstas [Quine (1960), p. 44]. Es de imaginar que podría existir algún argumento al respecto.

Lo que convence a Quine, sin embargo, no es el carácter convincente de estos ejemplos, sino su conocimiento de que deben existir tales casos, y éstos le parecen los mejores candidatos. ¿Cómo sabe que es así como son las cosas? La respuesta puede encontrarse en una afirmación que ya hemos citado: «las oraciones observacionales son aquellas que aprendemos a comprender en primer lugar, tanto como niños como en calidad de lingüistas de campo».

Es claramente posible aprender un lenguaje sin conocer ninguno con anterioridad, dado que todos hemos hecho tal cosa. Pero, si el holismo es verdadero y el significado de cada oración depende de los significados de las otras, ¿cómo lo hicimos? No parece que hay ningún posible punto de arranque, ninguna oración cuyo significado fuera, por decirlo así, lo bastante autosuficiente como para poder ser aprendida como el primer paso para el aprendizaje de las restantes. Pero debe haber algún lugar desde el que comenzar, algo que quien está aprendiendo, sea niño o lingüista, puede ponerse en las alforjas y utilizarlo como dato firme con el que evaluar hipótesis sobre lo que pudieran significar otras oraciones. Por lo menos, nuestro holismo debe atemperarse respecto a los requisitos impuestos por el aprendizaje del lenguaje. Debemos reservar algún lugar para el atomismo en la teoría del significado, dado que, en caso contrario, convertiríamos en algo imposible el aprendizaje lingüístico (radical).

El fundamentalismo de Quine se revela en el contraste entre el holismo al nivel no–observacional y el atomismo al nivel observacional. Las asimetrías a este respecto son evidentes. Las oraciones observacionales, en las que el significado está fijado y la traducción es posible, son la única puerta de entrada al lenguaje; ésta es la asimetría semántica. Las oraciones de observación pueden verificarse individualmente, nuestra aceptación de ellas puede justificarse una a una, y constituyen la evidencia sobre la que ha de descansar lo no-observacional, es decir, la totalidad de la ciencia; ésta es la asimetría epistemológica. Para Quine, hay datos y hay teoría, y, con independencia de las características internas positivas que pudiera tener la teoría, su justificación se logra, si es que se logra, como nos enseñó nuestra semántica verificacionista; es decir, por apelación a la diferencia que establecería la verdad de la teoría con relación a la experiencia posible y por verificación directa (fuerte) de si la experiencia se comporta como la teoría dice que debe hacerlo. Así pues, la posición de Quine muestra los rasgos características del fundamentalismo.

7.3. ATOMISMO Y HOLISMO

El atomismo de Quine en el nivel observacional se fundamenta en su adhesión al verificacionismo como expresión de empirismo en teoría de significado. Y considera que esto demuestra que todo buen empirista debe caminar junto a él. «El tipo de significado que es básico para la traducción, y para el aprendizaje del propio lenguaje, es

necesariamente el significado empírico y nada más [...]. Ciertamente, no nos queda otra opción más que la de ser empiristas en lo que respecta a la teoría del significado lingüístico» [Quine (1969), p. 81]. Éste es el pensamiento que proporciona el argumento crucial en favor del fundamentalismo y de sus asimetrías: el de que el empirismo conduce al verificacionismo y que el verificacionismo incorpora las asimetrías del fundamentalismo. Por lo que nadie puede ser empirista sin ser también fundamentalista.

No creo que esta conclusión sea correcta. Contra ella argumentaré, en primer lugar, que lo que Quine considera verificacionismo no es la única teoría empírica del significado y, en segundo lugar, que una teoría que abandone el atomismo al nivel observacional en favor de un holismo más completo sería preferible, en líneas generales, tanto para los empiristas como para los demás.

Comenzamos con tres principios expuestos en 6.2: PV, PC y PS. Se dijo que PC y PS se seguían «naturalmente» de PV, pero tal cosa no era toda la verdad. PV trata de las condiciones en las que una oración es significativa —en contraposición a que fuera un sinsentido— mientras que PS trata de las condiciones determinantes de que una oración tenga un significado más bien que otro. Deberíamos confiar en que una teoría del primer tipo nos ayudara con cuestiones del segundo tipo, pero no podríamos exigir por anticipado que sucediera tal cosa. Ciertamente, es compatible con PV, tal y como los hemos expresado, que dos oraciones tengan las mismas consecuencias observacionales aunque difieran en significado. Ésta es la posibilidad que prohíbe PS. Por tanto, es obvio que PS va mucho más allá de PV; quizá sea una de sus consecuencias naturales, pero no es su consecuencia «lógica».

El problema radica en el paso de PV a PC. PV todavía deja abierta la posibilidad de que algunas oraciones tengan, como componente de su significado, algo más que el mero efecto que establecería su verdad respecto a la evidencia de nuestros sentidos. Y, si existen oraciones de ese tipo, PC no será verdadero en su formulación actual. Habrá entonces oraciones cuyas consecuencias observacionales conocemos, sin que conozcamos otras posibles contribuciones (sean las que sean) a su significado. Pero si abandonamos PC, tal y como está formulado, ya no podemos llegar a PS. Una versión más correcta y más débil de PC,

PC^2: *a* conoce el significado de *p* → *a* conoce la diferencia que establecería la verdad de *p* respecto a las evidencias sensoriales de *a*,

no nos permite aseverar el tipo de identidad consagrada por PS, dado que ya no contiene el bicondicional «si y sólo si».

Si esto es así, Quine no tiene razón al afirmar que nadie puede ser empirista respecto a la teoría del significado sin aceptar su particular versión de verificacionismo, centrada en PS. Deriva el holismo de su teoría del significado de PS, pero insiste en que no podemos explicar el aprendizaje del lenguaje como actividad empírica, a menos que permitamos que algunas oraciones tengan un significado determinado y que, por tanto, sean independientemente verificables (de un modo fuerte).

El ataque a Quine tiene, por tanto, dos polos, el rechazo de PS y el intento de mostrar que no necesitamos ser atomistas para poder explicar la posibilidad del aprendizaje del lenguaje. El primero ya se ha alcanzado; Quine no tiene razón al vincular el empirismo con PS. Si PV basta para ser empirista, PS no es necesario.

¿Por qué no es posible suponer que los datos iniciales con los que comenzamos a aprender el lenguaje no son completamente sólidos, y que su papel es el de ser revisados, reevaluados y quizá abandonados a la luz de lo que suceda más tarde? Parece obvio que para comenzar sólo necesitamos oraciones cuyo significado esté vinculado estrechamente a lo que no es inmediatamente accesible, no que existan oraciones cuyo significado pueda darse completamente en términos de sus condiciones de asentimiento. Todo lo que se necesita es una diferencia de grado, no de tipo. Hay algo de cierto en la idea de que nadie podría comenzar a aprender el lenguaje partiendo de las leyes de la física y de la lógica; se necesitan muchas cosas antes de que podamos comenzar a entender éstas. Pero todo esto no muestra que las oraciones a partir de las que *podemos* empezar deban ser de un tipo distinto al de ellas; sólo muestra que su grado de observacionalidad debe ser mucho mayor. Si siempre hay un elemento del significado de nuestras oraciones que no es observacional, la mera observación no nos revelará la totalidad del significado. Pero esto sólo muestra que, en el proceso de aprendizaje, no podremos aprender todo el significado de una vez. No muestra que no podamos comenzar, sólo que nuestros movimientos iniciales serán reevaluados más tarde. Por tanto, no necesitamos abrazar la asimetría semántica para poder explicar nuestro aprendizaje del lenguaje. Y, en ese caso, la asimetría epistemológica que depende de ella se desmorona también. Parece, pues, que existe la posibilidad alternativa de adoptar un holismo más completo que incluya las oraciones de observación, sin dejar de ser coherentes con las exigencias del empirismo. Queda ahora por demostrar que es ése el camino correcto que hay que seguir.

7.4. LOS MÉRITOS DE UN HOLISMO MÁS COMPLETO

Buscamos razones para ser más quineanos que el propio Quine, para ir más allá de la indeterminación de lo que el propio Quine querría. ¿Bastaría con decir que las razones de Quine para la indeterminación en el nivel no–observacional debieran abarcar también lo observacional, y que su razón para salvar lo observacional no llega a ser contundente? Quizá Quine estuviera de acuerdo en que la distinción entre centro y periferia en la teoría siempre es una mera cuestión de grado, y que no podemos pasar abruptamente de la indeterminación a la determinación.

Existe cierta ironía en que esta oferta parezca un compromiso, dado que la filosofía de la ciencia de Quine insiste una y otra vez en que no hay una distinción firme entre observación y teoría. Si él mismo parece olvidar esta obsesión cuando se enfrenta a la teoría del significado, es por su convencimiento de que el aprendizaje del lenguaje requiere algún punto de determinación. Si abandonamos esta exigencia, podemos contentarnos con adscribir un grado menor de indeterminación a lo observacional. Ésta es la posición que parece haber sido adoptada por Quine en algunos escritos anteriores [Quine (1960), § 10, y Quine (1953), cap. 2].

Sean cuales sean los méritos de este compromiso, no es un compromiso completamente afortunado, dado lo que hemos dicho anteriormente. Quizá se dan en él tres dificultades distintas.

La primera es el hecho de que uno de los fundamentos del argumento de Quine para la indeterminación era su teoría empirista del significado, PS. Hemos visto que no hay ninguna razón para aceptar PS, y hemos considerado que PV, que es más débil, es suficiente para un empirista. De modo que no podemos seguir la ruta inicial de Quine para fundamentar la indeterminación.

La segunda dificultad consiste en que Quine, al insistir en que la unidad de significado empírico es la globalidad de la ciencia, debe conceder que la totalidad de la teoría no sufre de la indeterminación que afecta al significado de las oraciones individuales. Parece, pues, que, si la unidad de significado determinado es la totalidad de la ciencia, en este nivel debemos encontrar la determinación de la traducción. Pero me da la impresión de que, incluso aquí, Quine insistiría en la indeterminación de la traducción.

La tercera es que el argumento que pasa de la indeterminación del significado a la indeterminación de la traducción puede parecer poco convincente. ¿Por qué el hecho de que oración no tenga significado

por sí misma implica que debe tener más de una traducción correcta? Ciertamente, todavía podríamos encontrar, para una oración cuyo significado está indeterminado, otra cuyo significado indeterminado se ajustara exactamente al de la primera. Podría ser más que convincente establecer la indeterminación de la traducción de un modo directo y usar esa conclusión para consolidar argumentos anteriores sobre la indeterminación del significado. Por todo esto, es probable que encontremos argumentos consistentes en favor de un holismo más general. El mismo Quine ofrece dos de estos argumentos, y existe un tercero en el seno de la tradición quineana.

En primer lugar, está lo que denomina «el argumento desde arriba». Este argumento trata de establecer la indeterminación de la traducción directamente desde la infradeterminación de la teoría por los datos. Para empezar, es preciso distinguir infradeterminación de indeterminación. Podríamos defender que la teoría está infradeterminada por los datos en el sentido de que dos teorías rivales abarcan igualmente bien los mismos datos, pero insistir en que una de ellas es correcta y la otra errónea, a pesar de que no podamos decir cuál es la correcta. La indeterminación va mucho más allá de esta infradeterminación de la teoría por los datos; mantiene que, dadas dos teorías igualmente satisfactorias, no hay ningún hecho relativo a cuál sea correcta. Quine sostiene que la traducción está tanto infradeterminada como indeterminada, mientras que nuestra teoría física sólo está infradeterminada. (Adopta la segunda posición porque desea dar el máximo sentido posible al sentimiento de que nuestra ciencia física o algo semejante a ella es especialmente verdadera; no necesitamos tal supuesto en el caso de la traducción.) El argumento es que la indeterminación de la traducción se sigue de la infradeterminación de la teoría física, dado que el nativo que tratamos de traducir tiene una teoría física que está infradeterminada por sus evidencias, y nuestra traducción de su teoría, al ser una teoría sobre cuál pueda ser su teoría, también está infradeterminada. Quine, por tanto, afirma que en la traducción «el mismo viejo desajuste empírico, la vieja indeterminación entre teorías físicas, reaparece en segunda intensión» [Quine (1970), p. 179; cf. Quine (1975b), pp. 302-304]. Pero este argumento no parece convincente. ¿Por qué una doble dosis de infradeterminación produce indeterminación, si una única dosis no lo hace? Parece que Quine pasa por encima de esta cuestión en el fragmento citado, escribiendo «indeterminación» donde debiera haber escrito «infradeterminación». Por tanto, creo que el «argumento desde arriba», tal y como está formulado, no consigue su objetivo.

Un segundo argumento quineano, más convincente, está relacionado con los criterios que se usan de hecho en la construcción de traducciones. Si los examinamos, descubriremos que hay más de un criterio, y que los distintos criterios compiten entre sí. ¿Qué es lo que pedimos de un manual de traducción? Lo usamos para interactuar con la gente que estamos traduciendo, y para comprenderlos a ellos y a sus prácticas. Nos encontramos con que operan conjuntamente distintos criterios, sopesados distintamente de acuerdo con los propósitos generales del traductor. Un criterio del éxito de la traducción es que nuestras traducciones de las creencias de los nativos las muestren verdaderas en líneas generales. Consideramos que una traducción será implausible en la medida en que impute errores masivos. Pero otro criterio es que nuestra traducción debe atribuir a los nativos las creencias cuya posesión por su parte tenga sentido para nosotros. Estos dos criterios se denominan normalmente el principio de caridad y el principio de humanidad; esta última denominación se la debemos a Grandy, que argumenta que el mero principio de caridad es insuficiente [Grandy (1973)]. Y estos dos principios pueden entrar en conflicto. A veces, deberemos escoger entre atribuir a los nativos una falsa creencia cuya adquisición por su parte nos resulte inteligible o una creencia verdadera cuya adquisición no podamos entender. Por ejemplo, podríamos tener que escoger entre decir que creen que hay un espíritu en el cable o que creen que existe cierta carga eléctrica. Cuál sea la opción escogida dependerá de cómo sopesemos los dos principios, el de caridad y el de humanidad, lo que, a su vez, dependerá, en parte, de nuestros propósitos y, en parte, de nuestra predisposición a aceptar que los nativos son, en sus creencias, radicalmente distintos de nosotros. Si estamos dispuestos a aceptar tal cosa, como muchos antropólogos de campo (que, en ocasiones, parecen no sólo predispuestos, sino ansiosos por suponerlo), es probable que podamos escribir un manual de traducción que tenga la apariencia de mostrar cómo son las cosas. Pero siempre existirá otro manual igualmente bueno que no lo haga del mismo modo, o, al menos, que difiera en grado.

Existen otros criterios que sería razonable aplicar a los manuales de traducción. Por ejemplo, que las traducciones presentadas sean simples (fácilmente aprendibles), y que nos capaciten para predecir lo que hagan los nativos. Pero ya se ha dicho bastante para establecer lo esencial: que con criterios diversos y conflictivos debemos esperar que haya más de un posible manual que los satisfaga igualmente bien. Cada manual satisface los criterios de una buena traducción, por lo que «da el significado» de la oración nativa; pero cada manual ofrece

una oración diferente en el lenguaje «de casa» como traducción. Es fácil ver hasta qué punto está indeterminada la traducción.

Es posible distinguir entre dos tipos de principios que gobiernan cualquier investigación intelectual. En primer lugar, los principios de evidencia. Establecen las condiciones que, en caso de cumplirse por una teoría, son evidencia de que la teoría es verdadera. Por ejemplo, en ciencia, podríamos considerar que una teoría que proporcione una explicación convincente de los datos es la que tiene más posibilidades de ser verdadera. Hay, en segundo lugar, principios regulativos relacionados con el uso práctico de una teoría. En ciencia podríamos suponer que la simplicidad es una virtud; una teoría más simple sería más fácil de manipular y falsar, pero eso no es una razón para considerar que tiene más posibilidades de ser verdadera. (Hay dos puntos de vista a este respecto; también es posible sostener que simplicidad es evidencia de verdad y que, por tanto, el principio de simplicidad no es regulativo.) La sugerencia presente es la de que los principios de caridad y humanidad no son principios de evidencia. Nadie supone que nuestra traducción tiene más posibilidades de ser correcta cuantas más creencias verdaderas atribuya a los nativos. Tales principios son principios regulativos. Su fuerza deriva de los propósitos prácticos de la traducción que son muy variables. Cuantas más creencias de los nativos supongamos falsas, más difícil resultará la interacción social con ellos.

El argumento para la indeterminación de la traducción nos exige que, una vez que se hayan satisfecho los principios de una buena traducción, incluso aunque puedan sopesarse de formas diversas y verse igualmente bien satisfechos por manuales diferentes, debemos considerar que el significado es lo que está preservado por esos manuales alternativos. Esto es decir que son los principios regulativos de la traducción los que dan contenido a la noción de significado. «Preservar el significado» no es ningún principio adicional para traducir. Si fuera distinto en contenido a los otros principios, regulativos, se trataría de un principio de evidencia. Pero no tiene un contenido distinto al de los principios regulativos, se limita a resumir su efecto combinado. Por tanto, el significado es lo que preservan las buenas traducciones.

Este argumento para la indeterminación de la traducción no ofrece ningún ámbito en el que sea posible una distinción entre lo observacional y lo no–observacional que no sea de grado. Y nos ofrece un modo de desplazarnos desde la indeterminación de la traducción a la indeterminación del significado. Es una de las razones que podemos ofrecer en favor de un holismo más completo en teoría del significado.

El tercero es un argumento relativo a la relación entre creencia y significado. En la traducción radical nos encontramos, al menos, con dos tareas distintas: establecer lo que quieren decir las oraciones nativas y descubrir las creencias de los nativos. Si supiéramos el significado de las oraciones, no tendríamos demasiados problemas con las segundas. Si conociéramos las creencias, sería muy fácil llegar al primero. Pero necesitamos ambas cosas a la vez. ¿Cómo no es posible comenzar? Debemos suponer que sabemos lo que creen en algún conjunto inicial de circunstancias para poder establecer hipótesis sobre lo que significan sus oraciones. Y éste es un supuesto que no podemos justificar en modo alguno. Todo lo que podemos hacer es pasar continuamente de una a otra parte, asumiendo las hipótesis sobre una de las vertientes como si fueran sólidas para poder poner a prueba las hipótesis sobre la otra vertiente, y, después, asumir estas últimas para poder poner a prueba las hipótesis que las han generado. La sugerencia es la de que la indeterminación surge, precisamente, de esta continua interacción. Un buen ejemplo es el de alguien que diga que tiene dos rinocerontes en la nevera y que exprime uno para el zumo del desayuno. ¿Qué sucede? ¿Quiere decir «naranjas» cuando habla de «rinocerontes»? ¿O utiliza las palabras con el mismo significado que nosotros y tiene esas extrañas creencias sobre los rinocerontes? Hay varios defectos en estas preguntas. La finalidad del ejemplo es, en primer lugar, la de mostrar que siempre existirán modos alternativos de contrapesar las atribuciones de creencia y significado cuando tratemos de comprender a los nativos y, en segundo lugar, cuestionar el sentimiento de que debe haber un hecho determinado al respecto. En el ejemplo anterior, ¿qué diferencia establecería para la persona en cuestión la verdad de una de las hipótesis alternativas? Para ella, no parece haber demasiada diferencia. Pero, si no hay diferencia, no puede estar determinada la cuestión sobre cuál de las dos hipótesis es la correcta. Las respuestas difieren, no porque exista un hecho determinado que marque la discrepancia relevante, sino porque son claves distintas para una explicación más general sobre la persona en cuestión.

Por tanto, esta interacción entre creencia y significado tiene la consecuencia de que la traducción está indeterminada. También, que la creencia está indeterminada; en el ejemplo anterior se sugirió que no había un hecho determinado relativo a lo que se cree, más de lo que pueda haberlo con relación al significado. El significado está indeterminado a causa de la indeterminación de la traducción.

Este último argumento no ofrece ningún ámbito en el que sea posible practicar la distinción entre lo observacional y lo no observacional.

Si no coincidimos con relación a la presencia de la rojez, nuestra discrepancia puede deberse tanto a una diferencia en nuestra creencia como a una diferencia en el significado. Tenemos, pues, dos argumentos que sugieren que nuestro holismo debe ser más completo. Creo que una teoría holista del significado no incorporaría las asimetrías características del fundamentalismo. Si la asimetría epistemológica de Quine surge de la asimetría semántica, el holismo en la teoría del significado es letal para ambas. En ese caso, tenemos una razón general para preferir una teoría no-fundamentalista, si es que la podemos encontrar. Una teoría holista del significado debe llevarnos a una epistemología holista. En el capítulo próximo, empezaré considerando la base de esa epistemología, la teoría de la coherencia sobre la justificación.

7.5. VERIFICACIONISMO, ANTIRREALISMO Y FUNDAMENTALISMO

El fenomenalismo clásico es una forma extrema de empirismo, es antirrealista. Tiene la esperanza de deshacerse del argumento del error mostrando que todos los enunciados verdaderos son, en principio, verificables (y, si fuera posible, no sólo débilmente verificables). Las formas menos extremas de empirismo son realistas. Quine mantiene que la teoría está infradeterminada por los datos; ésta es una posición realista. El fundamentalismo puede expresar cualquiera de estas dos formas de empirismo; en tanto que teorías del significado, contienen las asimetrías que caracterizan la epistemología fundamentalista. Pero si el fundamentalismo se desprende de esas formas, más fuertes o más débiles, de verificacionismo y acepta una dosis mayor de holismo en teoría del significado, se produce una tensión sintomática en su nueva posición: el contraste entre su holismo en teoría del significado y su atomismo en epistemología. La distancia entre ambos quiere decir que el fundamentalismo debe apoyarse ahora en los dos argumentos «internos» sobre el regreso de justificaciones, y no en otros supuestos filosóficos (véase 4.1). Si estos argumentos se vinieran abajo, o fueran inaplicables en el seno de una posición epistemológica más bolista, se encontraría de nuevo sin ningún apoyo. Argumentaré que es esto lo que sucede. Mi conclusión será que, en ausencia de razones convincentes para admitir las asimetrías características del fundamentalismo, será mejor que adoptemos una teoría que esté libre de ellas.

LECTURAS COMPLEMENTARIAS

Véanse Quine (1960, cap. 2), sobre la traducción radical, y Quine (1970), sobre los argumentos respecto a la indeterminación de la traducción.

Hookway (1978) discute esos argumentos de un modo constructivo.

Dummett (1978, cap. 2) arguye que la tesis de la indeterminación de la traducción es o bien verdadera y banal, o bien significativa pero errónea.

Una cuestión importante en la filosofía de Quine es hasta qué punto la tesis de la indeterminación de la traducción es independiente de la tesis de la indeterminación del significado oracional (6.3) y, por tanto, también independiente del rechazo de la distinción analítico/sintético. Dummett (1978, cap. 22) y Hylton (1982) lo discuten.

Davidson (1974) discute problemas asociados con la interacción entre creencia y significado; también Lewis (1974).

8. TEORÍAS DE LA COHERENCIA

8.1. ¿QUÉ ES LA COHERENCIA?

En los dos últimos capítulos comenzamos a tratar nuestras creencias como un tipo de teoría interrelacionada, y nos planteábamos el problema de cómo están relacionadas. Por supuesto, hay muchos aspectos de esta cuestión que no hemos examinado todavía, aunque sí hemos encontrado razones para rechazar una posible respuesta. Se trata del punto de vista de que la relación es crucialmente asimétrica, de que hay una distinción asimétrica entre evidencia y teoría, según la cual la evidencia confirma y desautoriza la teoría de una manera en la que la teoría no puede confirmar o desautorizar la evidencia. El fundamentalismo nos ofrece ese tipo de estructura cuando asevera que la justificación es unidireccional, y cuando pretende que hay algunos puntos en la estructura, las creencias básicas, que son relativamente fijos. La noción de inferencia desde puntos fijos involucra claramente las asimetrías relevantes. La misma noción de inferencia es asimétrica. Es posible inferir B a partir de A sin poder inferir A a partir de B.

La noción de coherencia, en la que se basa una teoría más completamente holista, pretende ser simétrica. Pero, para conocer el éxito de tal pretensión, necesitamos saber con mayor exactitud qué quieren decir los coherentistas por «coherente».

Todos los coherentistas están de acuerdo en que la consistencia es una condición necesaria para la coherencia. Bradley añadió (1914, pp. 202-203) que un conjunto coherente debe ser completo o comprehensivo en algún sentido. (Enseguida veremos por qué.) Pero la consistencia y la compleción no son suficientes, no hacen justicia a la intuición de que el conjunto coherente se mantiene unido y mutuamente ajustado de un modo particular. Para hacer justicia a ello, los coherentistas clásicos usan la noción de implicación (*p* implica *q* si y sólo si, dada *p*, *q debe* ser verdadera). Brand Blanshard escribió que en un sistema completamente coherente «ninguna proposición sería arbitraria, cualquier proposición estaría implicada, conjunta o individualmente, por las otras, ninguna proposición podría mantenerse fuera del siste-

ma» [Blanshard (1939), vol. 2, pp. 265-266]. Pero este análisis de la coherencia en términos de implicación mutua es discutible. Ewing sugiere que bastaría con que cada uno de los miembros de un conjunto coherente estuviera implicado por el resto [Ewing (1934), p. 229] y que todo lo que fuera más allá de eso resultaría desastroso. En realidad, ¿podemos dotar de sentido a la idea de un sistema cada uno de cuyos miembros entrañe a todos los demás?

En lugar de responder de un modo directo a esta pregunta, podemos encararla si consideramos una objeción a cualquier uso de la noción de implicación mutua como el elemento central de un conjunto coherente. Esa noción, tal y como la usa Blanshard, es suficientemente simétrica. Pero la implicación entendida al modo tradicional no es asunto de grados. Lo que es importante, porque los coherentistas pretenden dar sentido a la idea de que, a medida que crece el propio conjunto de creencias, mejora (eso esperamos) y gana en coherencia. Y no sólo porque se convierta en algo más completo; difícilmente la compleción puede considerarse como una virtud en sí misma. Es importante percibir que no podemos apoyarnos en la consideración de que las relaciones de implicación sólo se mantienen entre los miembros de un conjunto completo, porque con tal cosa no atraparíamos realmente el sentido en el que, al ampliar nuestro conjunto de creencias, tratamos de hacerlo más coherente. Dado que es probable que nunca logremos llegar a un conjunto coherente y completo, la definición de coherencia en términos de implicación tiene la consecuencia de que no hay nadie cuyas creencias sean coherentes en absoluto. [Más problemas en la apelación a la implicación se exploran en Rescher(1973), cap. 2.5.]

De modo que, para tener una teoría de la justificación como coherencia, deberemos dar sentido a la idea de que la justificación puede crecer. Un análisis alternativo de la coherencia, debido a Lehrer (1974) y Sellars (1973), define un conjunto coherente como uno que es consistente, completo y mutuamente explicativo. La idea aquí será que, a medida que el conjunto aumenta de tamaño, podemos esperar que cada uno de sus miembros sea explicado mejor por el resto. Las explicaciones pueden mejorar en calidad; esto da cuenta de la mejora en la justificación. Y la noción de explicación *mutua* es claramente simétrica, en el sentido requerido.

Podrían hacerse dos comentarios sobre este análisis de lo coherente como lo mutuamente explicativo. Parece, en primer lugar, que el requisito de la compleción puede ser dejado a un lado. La exigencia de la compleción puede fundamentarse en la búsqueda de un grado

más alto de explicatividad mutua. Lo que está muy bien, dado que en realidad no disponemos de una noción clara de compleción. Quizás pudiéramos suponer que un conjunto completo contiene toda proposición o su contradictoria. Pero esto no nos servirá de ayuda, a menos que tengamos una idea clara de qué es «toda proposición». Del mismo modo, tampoco poseemos una idea clara de una explicación perfecta, de alcanzar una situación desde la que ya no sea posible mejorar. Esta situación, como mucho, será un punto límite de nuestras constantes aproximaciones. Pero no hay ningún contenido en el supuesto de que la hemos alcanzado. Por todo ello, las dudas sobre la compleción nos permiten dejarla tranquilamente a un lado al definir la coherencia. (Más tarde, surgirán otros motivos.)

El segundo comentario es que la coherencia es una propiedad de un conjunto de creencias, no de sus miembros individuales. El conjunto es coherente en la medida en que sus miembros sean mutuamente explicativos y consistentes. Lo que será importante para las reflexiones siguientes.

Puede parecer que este análisis en términos de explicación mutua es una mejora del que apela a la implicación para mantener unido al conjunto coherente. Pero creo que el análisis de la explicación mutua restituye más que reemplaza la utilización que hacía Blanshard de la implicación. La manera en que entiende Blanshard la implicación no es la tradicional. En la interpretación tradicional, «*p* implica *q*» se entiende de un modo atomista, como un rasgo de los significados individuales de *p* y de *q*; dado el significado de p y el de *q*, si *p* es verdadero debe serlo *q*, con total independencia de cualquier consideración sobre otros rasgos del sistema. En esta interpretación de la implicación se basa Rescher cuando se queja de que, cuando *p* implica *q*, *q* es un miembro redundante del conjunto; de modo que un conjunto coherente está infectado de redundancia mutua, contrariamente a la intención explícita de Blanshard. Lo que también está en la base de la observación anterior de que la implicación no es asunto de grados. Pero Blanshard, como es de esperar en alguien que es holista en teoría del significado, no concibe la implicación de ese modo. Para él, la implicación sólo se da en el seno de un sistema; y, dado que el sistema determina los significados de *p* y de *q*, determina la intensidad del vínculo entre *p* y *q*. Por lo que el vínculo puede hacerse más estrecho a medida que el sistema crece.

Hay siempre un obvio vínculo intuitivo entre implicación, como la considera Blanshard, y explicación. Explicar *q* apelando a *p* es mostrar por qué, dada *p*, *q* debe ser verdadera. La explicación funciona en la

medida en que muestra que, dada *p*, *q* debe ser verdadera. De modo que, en el sentido de Blanshard, la explicación revela la implicación. Y, como la implicación, la explicación debe ser considerada desde el punto de vista holista más bien que de un modo atomista. Por tanto, al final, los dos análisis de la coherencia acaban coincidiendo.

Antes de dirigir nuestra atención a la teoría de la justificación como coherencia debemos considerar la teoría de la verdad como coherencia; las dos están estrechamente vinculadas.

8.2. LA TEORÍA DE LA VERDAD COMO COHERENCIA

Esta teoría mantiene que una proposición es verdadera si y sólo si es un miembro de un conjunto coherente.

Quien mantenga dudas sobre la posibilidad de un conjunto completamente coherente mantendrá, con los defensores clásicos de la coherencia, que la verdad es asunto de grados. Las proposiciones son verdaderas en la medida en que hay un conjunto coherente del que son miembros. Debemos advertir, sin embargo, que la teoría no identifica verdad con coherencia. No dota de sentido a la noción de un conjunto verdadero. Por el contrario, define la verdad para los miembros individuales del conjunto. Una proposición es verdadera si y sólo si es un miembro de un conjunto coherente. Las proposiciones no pueden ser coherentes en el sentido requerido, y los conjuntos sólo pueden ser calificados como verdaderos en la medida en que son miembros de conjuntos más amplios.

Sin embargo, la teoría sí pretende proporcionar una definición de verdad. No se limita a decirnos en qué circunstancias estaríamos justificados al considerar que una proposición es verdadera. Podría hacer tal cosa afirmando que estamos justificados a creer que/? es verdadera en la medida en que tal supuesto incremente la coherencia de nuestro conjunto de creencias. El coherentista tiene esa pretensión: ofrece un análisis *criterial* de la verdad, una teoría sobre cuáles sean los criterios para la verdad. Pero también nos ofrece un análisis *definitional*. Se supone que, como veremos, los dos análisis encajan entre sí.

Muchos filósofos que han mostrado cierto interés en la teoría de la verdad como coherencia han discutido que la teoría ofrezca una definición de verdad, argumentando que, considerada de esa forma, la teoría es manifiestamente falsa [por ejemplo, Russell (1907)]. Es manifiestamente falsa porque, independientemente de cuánto ajustemos el análisis de la coherencia, tendremos que admitir que siempre hay más

de un conjunto coherente de proposiciones. Nada en la noción de coherencia, tal y como está definida, nos da derecho alguno a decir que hay un único conjunto con la máxima coherencia. Pero, obviamente, como mucho sólo puede haber un conjunto completo de verdades. De modo que la verdad no puede definirse sólo en términos de la coherencia.

La situación puede recordarnos las tesis de Quine sobre la infradeterminación de la teoría por la evidencia. Si hay más de una teoría igualmente efectiva para hacer frente a la experiencia, ¿qué debemos decir de diferentes teorías? ¿Sería posible afirmar que son todas verdades, o que todos sus miembros son verdaderos? Parece que no podemos. Si nuestros distintos conjuntos coherentes están todos cerca de la compleción, si constituyen diferentes descripciones del mundo igualmente completas, ¿cómo admitir que todas las partes de esas descripciones diferentes son igualmente verdaderas? Ciertamente, si las descripciones son diferentes, están en mutua competencia, y el premio por el que compiten no es otro que el de la *verdad*. Por lo que sólo uno de los conjuntos competidores puede contener exclusivamente verdades, y la teoría de la verdad como coherencia es incorrecta.

Ésta es una objeción habitual a la teoría de la verdad como coherencia. Podemos denominarla la objeción de la pluralidad. Produce una enorme indignación entre los que se clasifican a sí mismos como coherentistas. Brand Blanshard escribe (1939, vol. 2, pp. 275-276):

> Esta objeción, como muchas otras críticas demoledoras, hubiera tenido más sentido si alguien hubiera mantenido alguna vez la teoría que se pretendía aniquilar. Pero es una enorme confusión en la medida en que pretenda representar el uso sensato de la teoría de la coherencia.

Blanshard argumenta que la objeción de la pluralidad no aprecia el carácter empirista de este coherentismo. De hecho, al igual que hacen otros coherentistas como Bradley, considera que sólo existe un conjunto coherente, y que tal conjunto se distingue de sus rivales por ser el único que tiene fundamento en la experiencia. Sucede así, de acuerdo con Bradley, a causa del mismo objetivo de la investigación y el pensamiento, que no es sino el descubrimiento de la ordenación más sistemática de nuestra experiencia [Bradley (1914), p. 210; cf. p. 203]:

> Mi experiencia es sólida [...] en la medida en que, en pocas palabras, constituye un sistema. Mi objetivo es el de tener un mundo tan comprehensivo y coherente como sea posible y, para poder alcanzarlo, no sólo tengo que reflexionar, sino que necesito el recurso constante a los materiales de los

> sentidos. Debo acudir a este fuente tanto para verificar viejos descubrimientos como para incrementarlos con lo que haya de novedoso. Y para ello debo depender de los juicios perceptivos.

Estos coherentistas afirman que la empresa debe comenzar con los datos de la experiencia para construir un conjunto de creencias sobre esos datos que los ordenará de la manera más sistemática (coherente). Para ello es posible que necesitemos rechazar *algunos* de esos datos, pero no podemos rechazarlos todos porque nuestro auténtico objetivo es dar sentido a lo que no es dado. De modo que el conjunto de creencias que construimos *debe* estar fundamentado en la experiencia, y este fundamento en los datos de la experiencia garantiza que sólo habrá un conjunto que constituya la «ordenación más sistemática».

Esta apelación a la necesidad de un fundamento empírico se las arregla para excluir todos los conjuntos artificiosos de proposiciones que podrían parecer coherentes. De modo que, por ejemplo, una expansión perfecta de las historias de Sherlock Holmes no debería contar como una descripción verdadera del mundo, a pesar de su coherencia. Pero, desgraciadamente, aunque excluyamos todos esos conjuntos, todavía nos quedará más de uno, dado que nada en la apelación a la necesidad de ordenar los datos de experiencia puede convertir en necesaria la existencia de una única ordenación máximamente sistemática. Es posible, a pesar de todo lo que hagamos por evitarlo, que haya más de una manera igualmente buena de «ordenan) los datos o de encajarlos en un sistema explicativo, particularmente cuando recordamos que algunos de los datos serán rechazados en el camino. Esto es, después de todo, la infradeterminación de la teoría por la evidencia. Con lo que la objeción de la pluralidad se mantiene en pie.

La defensa correcta contra la objeción de la pluralidad es el ataque. Deberíamos preguntar si existe alguna otra teoría de la verdad, algún otro análisis de lo que sea la verdad, que se comporte mejor. Enseguida parece que ninguna de las teorías habituales de la verdad tiene la consecuencia deseada de que sólo pueda haber un conjunto de verdades. Ciertamente, el adversario tradicional de la teoría de la coherencia, la teoría de la correspondencia, se enfrenta al mismo problema. Las teorías de la correspondencia tratan de erigir su análisis de la verdad sobre la observación inneglable de que la verdad de una proposición consiste en su ajuste con los hechos. Pero, en la medida en que hechos y proposiciones se mantengan separados entre sí, ¿qué puede evitar la existencia de dos conjuntos de proposiciones que se «ajusten a los hechos» igualmente bien? O bien debemos admitir que

la objeción de la pluralidad es tan efectiva contra una teoría como contra otra, y abandonar la exigencia de que la «verdad» es de algún modo única, o bien admitir que, aunque la verdad sea única, el mostrarlo no es misión de la teoría de la verdad.

Es posible que la objeción de la pluralidad tenga algún fundamento. Después de todo, el coherentismo debe admitir que las teorías alternativas son verdaderas en la medida en que sean igualmente coherentes, mientras que el teórico de la correspondencia puede decir que una es verdadera y la otra falsa. El teórico de la correspondencia tiene esta ventaja porque dice que el mundo es algo distinto a las teorías alternativas, que está más allá de ellas y que puede hacer que una sea verdadera y la otra falsa. Parece, pues, que el coherentista no puede dar un buen sentido a la idea de que diferentes teorías compiten entre sí o son incompatibles. Es ésta una debilidad que no afecta a su oponente.

La respuesta a esto tiene dos aspectos. En primer lugar, debemos decir que, para el coherentista, una teoría es incompatible con otra cuando no podemos aceptar ambas a la vez, so pena de perder coherencia. Desde el punto de vista de alguien con una teoría, cualquier otra teoría es falsa porque no puede añadirse a la teoría verdadera. Y, en segundo lugar, es sólo desde el punto de vista del mundo, un punto de vista externo a cualquier teoría, desde donde adquiere ventaja el teórico de la correspondencia. Sólo aquellos que no sostienen ninguna teoría en absoluto y que consideran todas las teorías desde un punto de vista externo pueden dar sentido a una noción de incompatibilidad entre teorías que vaya más allá de la propuesta coherentista. Pero no hay un punto de vista externo y libre de teoría. De modo que el coherentista puede ofrecer un análisis de la incompatibilidad de dos teorías coherentes, sin que exista otro análisis alternativo que sea propio de su adversario.

8.3. LA TEORÍA DE LA JUSTIFICACIÓN COMO COHERENCIA

Esta teoría sostiene que una creencia está justificada en la medida en que el conjunto de creencias al que pertenece es coherente. Cada creencia debe evaluarse apelando al papel que desempeña en el conjunto de creencias. Si el abandono de una creencia y, a veces, su sustitución por la creencia opuesta aumentara la coherencia del conjunto, tal creencia no estaría justificada. Si el conjunto es más coherente con esa creencia como miembro más bien que con una creencia alternativa, la creencia está justificada. La noción de justificación es relativa a

los creyentes individuales. El análisis completo debe ser: si el conjunto de creencias de *a* es más coherente con la creencia de que *p* como miembro que sin ella o con otra alternativa, *a* está (o estaría) justificado para creer que *p.*

¿Cuál es el vínculo entre justificación y verdad? Un conjunto de creencias con un grado razonable de coherencia justificará a cada uno de sus miembros. Pero ello no significa que todos sean verdaderos. Es posible que el conjunto no pueda ser expandido; que, más allá de cierto punto, cualquier adición de creencia no haga más que disminuir la coherencia del todo que va creciendo. En ese caso, los miembros del conjunto no pueden ser todos verdaderos, porque no pueden ser todos miembros de una totalidad auténticamente coherente. Pero todavía están justificados (para *a*). Igualmente una creencia puede ser verdadera cuando la proposición que es su contenido pertenezca a un conjunto coherente, sin que ello implique que está justificada para *a.* Según el análisis coherentista, por tanto, una creencia puede ser verdadera sin que esté justificada y estar justificada sin ser verdadera. La justificación puede crecer, pero no es preciso que con su crecimiento se aproxime a la verdad. Es obvio que, a medida que un conjunto de creencias crece y se convierte en algo más coherente, tenemos más razones para suponer que sus miembros son verdaderos. Pero es posible que no lo sean; de hecho, siempre será bastante probable que expansiones posteriores requieran que revisemos algunas creencias.

A pesar de la distinción entre creencia y justificación, el coherentista subraya como una virtud de su teoría el que, según ella, la verdad y la justificación vayan a la par. Los miembros de un conjunto de creencias van a estar justificados por la coherencia de éste; la coherencia de un conjunto de proposiciones, creídas o no, va a convertir a sus miembros en verdaderos. El sentido en el que, de acuerdo con la teoría, la verdad no coincide con la justificación no disminuye la ventaja de tener un vínculo confortable entre verdad y justificación.

Supongamos que, como sugieren Ewing (1934), Rescher (1973) y Lehrer (1974), adoptamos una teoría de la justificación como coherencia, mientras rechazamos la teoría de la coherencia de la verdad —es posible que estemos impresionados por la objeción de la pluralidad—. Nos enfrentamos a un misterio: ciertamente, nuestra teoría deberá mostrar de algún modo por qué es valiosa la justificación, por qué deben buscarse y aceptarse las creencias justificadas y rechazarse las que no lo están. Es obvio que mostraríamos tal cosa si mostráramos que, o cómo, las creencias justificadas tienen más posibilidades de ser verdaderas. Si consideramos que la coherencia es tanto un crite-

rio de verdad como de justificación, tenemos gran probabilidad de conseguirlo. La alternativa es suponer que la justificación es asunto de coherencia interna, una cuestión de ajuste entre objetos que son del mismo tipo, mientras que la verdad es asunto de correspondencia entre proposiciones y objetos de un tipo distinto, hechos o estados de cosas. Pero, en ese caso, sería difícil encontrar una razón para pensar que, cuando se da la relación interna de justificación, también se da la relación externa de verdad. De modo que nos encontramos con una ventaja enorme al poseer teorías de la verdad y de la justificación que se ajustan entre sí. La teoría de la verdad debe ajustarse a la epistemología y no debemos permitir que vaya al margen de ella.

¿Qué objetos están vinculados por la relación de explicación mutua en un conjunto coherente? En la teoría de la verdad como coherencia son proposiciones; en la teoría de la justificación como coherencia, también son proposiciones. De modo que, cuando hablamos de la justificación de la creencia de *a* de que *p*, estamos preguntándonos si la proposición *p* forma, con otras proposiciones creídas por *a*, un conjunto que promete ser coherente. Si lo hace, la verdad de *p* se explica por relación a la verdad de esas otras. [Una perspectiva diferente la adopta Lehrer (1974) al sugerir que lo que necesita explicación no es tanto la verdad de *p* como el hecho de que *a* crea que *p*.] Por tanto, también en este extremo, nuestra teoría de la verdad se ajusta a la teoría de la justificación.

Por otra parte, el coherentista diría que tenemos motivos más importantes para deshacernos de las asimetrías del fundamentalismo. Todavía no hemos descubierto ninguna razón convincente para adoptar esas asimetrías (véanse, sin embargo, 8.4 y 5.1) y, en ausencia de tal razón, debemos considerar que sólo hay una forma de justificación: la misma para todas las creencias. No existe ningún otro punto estable por apelación al cual puedan evaluarse las otras creencias. Todas las creencias se evalúan del mismo modo, teniendo en cuenta sus efectos en la coherencia global, por lo que no hay restricciones en lo que puede tomarse en consideración para dar apoyo a una creencia. La prueba crucial, como dice Bradley, es el sistema y no un criterio unidireccional de adecuación a la evidencia. Del mismo modo, cuando surge una dificultad no hay requisitos previos acerca de dónde debe efectuarse la revisión. No tenemos razones independientes para preferir retener creencias altamente observacionales antes que otras teóricas. La revisión correcta es la que resulta en una nueva totalidad más coherente, sin que podamos decir por anticipado qué tipo de revisión tiene más posibilidades de éxito.

Los coherentistas asegurarían que esta teoría holista se adecua a nuestra práctica real mucho mejor que la explicación fundamentalista, que es más restrictiva. En la práctica, no existen tabúes sobre qué puede utilizarse para dar apoyo a una creencia ni exigencias sobre los tipos de enunciados que deben retenerse preferentemente en caso de conflicto con otros. No siempre preservamos lo observacional a expensas de lo teórico. Por ejemplo, suponemos que alguien está equivocado cuando dice que la cortina le parece naranja, sobre la base de que los objetos con determinada estructura molecular no parecen de color naranja. Del mismo modo, defendemos nuestras creencias observacionales apelando a las teóricas (por supuesto —véase 4.3—, una forma débil de fundamentalismo lo podría admitir también). No sólo no hay una necesidad teórica de aceptar las asimetrías, sino que nuestra práctica revela que no hacemos tal cosa.

Es así como el coherentismo hace de la necesidad virtud. En ausencia de puntos fijos y de claves respecto a dónde comenzar la revisión, sabemos que, en cualquier momento, nuestro conjunto de creencias es meramente provisional. La necesidad de revisión puede aparecer por cualquier lado. Es ésta una forma de falibilismo (véase 4.2). Los coherentistas la aceptan con los brazos abiertos y aseguran que es su punto de vista el que revela la fuerza del falibilismo. El falibilismo no es un defecto infortunado, sino una parte esencial de la tarea epistemológica, la disposición a la revisión constante a la búsqueda de mayor coherencia.

La capacidad de la teoría para justificar los principios de inferencia que utilizamos le proporciona un soporte adicional. El fundamentalismo supone que necesitamos no sólo creencias básicas, sino también principios de inferencia que nos lleven de esas inferencias a la supraestructura más sofisticada. Es posible que comprendamos qué justifica las creencias básicas, pero ¿qué justifica los principios de inferencia? El ejemplo clásico de este planteamiento es nuestro tercer argumento escéptico sobre la inducción (1.2). Volveremos sobre ello en el capítulo 13. Pero el principio inductivo no es el único principio de inferencia que está en cuestión. Otros podrían ser los siguientes:

1. Si me parece recordar que hice una acción, es probable que la hiciera. (Memoria.)
2. Si los demás me dicen que observaron un suceso, es probable que el suceso ocurriera. (Testimonio de los demás.)
3. Si me parece que cierto objeto está frente a mí, es probable que esté frente a mí. (Percepción.)

Tales principios no pueden justificarse apelando a las creencias básicas, ni como conclusiones de inferencias a partir de esas creencias. Parece, por tanto, que el fundamentalismo todavía debe encontrar una forma ulterior de justificación para sus principios de inferencia. Los coherentistas se enfrentan a una tarea mucho más fácil, tanto si pueden alcanzar el éxito como si no. Para ellos, los principios de inferencia son, por supuesto, necesarios como una de las formas de mantener unido al conjunto coherente. Pero pueden justificarse del modo ahora familiar, consistente en la apelación al incremento de la coherencia que resulta de la adopción de un principio. Es de esperar que el uso de un principio incremente el tamaño de un conjunto de creencias, y que esté justificado si el conjunto aumenta en coherencia a la vez que en tamaño. Y si se dan principios en conflicto, como cuando consideremos una opción a 1 que incluye una restricción a ciertas circunstancias, está justificada la alternativa cuyo uso incremente más la coherencia del conjunto. (Estas cuestiones se examinarán mejor en los capítulos 11-13, donde se pondrá en cuestión esta respuesta.)

Otra ventaja del coherentismo, sugerida por Rescher (1973, p. 332), es la de que no presta atención exclusiva a la lucha del individuo por construir su propia epistemología, que es la concepción clásica de la tarea epistemológica (véase 5.1); en lugar de ello, da sentido a la idea de conocimiento como un fenómeno social, algo que puede compartirse y que se incrementa por el hecho de ser compartido. Esta pretensión parece depender de la facilidad con la que los coherentistas puedan justificar el uso del principio 2. Es como si los coherentistas comenzaran con el problema egocéntrico tradicional respecto a cuáles de nuestras creencias están justificadas. En este respecto no se alejan de la tradición, excepto en que no insisten en restringir los datos iniciales a hechos básicos sobre los propios estados sensoriales. El testimonio de los demás puede usarse de un modo más o menos inmediato para incrementar la coherencia del propio conjunto de creencias, por lo que es posible realizar un pronto desplazamiento desde el predicamento egocéntrico y pensar en uno mismo como en un colaborador, incluso como alguien que tiene más posibilidades de aprender de los demás que de contribuir por sí mismo a la suma total del conocimiento (una especie de modestia epistemológica). Esto no llega a ser el supuesto de que el conocimiento es un fenómeno completamente social, como algunos podrían desear, pero se aproxima a esa posición, a pesar de comenzar en el punto de partida tradicional.

Los coherentistas también suponen que su punto de vista, así como proporciona una posible justificación de la inducción, ofrece también

una perspectiva desde la que es posible eludir, si no refutar, al escéptico. Consideramos esta pretensión en 9.3, 9.5 y 13.3.

Por último, ya vimos en el capítulo 7 una razón general para buscar un holismo más completo en teoría de la justificación, para adecuarse a la teoría holista del significado. El coherentismo es *la* teoría holista; nos proporciona lo que buscábamos.

Son éstas las ventajas principales que los coherentistas reclamarían para su teoría. Ahora consideraremos el ataque crucial contra el coherentismo. Se trata de la queja de que no es compatible con el empirismo.

8.4. EL PAPEL DE LOS DATOS EMPÍRICOS

En 8.2 consideraremos la objeción de la pluralidad a la teoría de la coherencia de la verdad y mencionamos una respuesta habitual. Se trata de que un conjunto coherente se selecciona entre otros muchos porque está fundamentado empíricamente. La tarea del pensamiento es la de comenzar a partir de los datos de la experiencia y construir un conjunto de creencias alrededor de esos datos que los ordene de la forma más sistemática. Pero esta perspectiva empirista parece revelar una dificultad para el coherentismo como teoría de la justificación, dado que parece reintroducir una distinción entre dos tipos de justificación. En este punto el coherentista debe ser *monista*; debe afirmar que la justificación es siempre del mismo tipo. Sin embargo, ¿no introduce la noción de *datos* empíricos una suerte de pluralismo? Una teoría exclusivamente interesada en la relación interna entre creencias parece incapaz de capturar el papel de los datos. Un dato es lo que es gracias a su fuente, no gracias a su relación con otras creencias. Sus pretensiones de que debe ser aceptado derivan de que sea parte de nuestro *input*, parte de lo que nos está dando la experiencia.

Ciertamente, debemos reservar algún lugar a la idea de que las creencias de alguien estén justificadas, al menos en parte, por referencia a algo que esté más allá de las creencias mismas; por referencia, de hecho, a su experiencia. Pero esto equivale a abandonar el monismo de nuestro coherentismo y volver al tipo de asimetrías características del fundamentalismo. Sólo el fundamentalismo puede proporcionar a la evidencia sensorial el papel especial que debe tener en cualquier explicación empirista de la justificación de la experiencia. El empirismo y el coherentismo son incompatibles.

Debemos preguntarnos, en primer lugar, por qué el coherentista

debe sentirse preocupado por este ataque. Se le ha puesto en la situación de mantener que conjuntos de creencias que no tienen relación alguna con la experiencia de nadie pueden tener todos los rasgos definitorios de la coherencia. Pero sólo concedería tal cosa si aceptara la distinción entre creencia y experiencia; y es ésta una distinción en la que no coincidirían todas las partes de esta disputa. Podemos construir una forma de coherentismo inmune al argumento anterior si mantenemos, como Kant, la imposibilidad de establecer una distinción adecuada entre los elementos «cognitivos» y sensoriales de la experiencia sensorial, o si mantenemos que toda experiencia es una forma de conocimiento o juicio (esto es, de adquisición de creencia) más bien que una forma de sensación. Si el coherentista exige, para que se dé la justificación, que todos los elementos cognitivos estén interconectados, no hay ninguna posibilidad, una vez que consideramos que la experiencia es cognitiva, de que creencias completamente desconectadas de la experiencia sensorial pudieran contar como justificadas.

Esta línea de defensa, por importante que pueda ser, no es completa. Incluso si aceptamos que la experiencia es una forma de creencia, todavía podemos insistir en una distinción entre creencias sensoriales y otras (sin poder especificar exactamente cómo ha de ser trazada), y volver a expresar, por medio de esa distinción, la demanda del empirista como la exigencia de que las creencias sensoriales den apoyo a las demás. Y es ésta una exigencia de algo que está más allá de la mera coherencia, dado que la noción relevante de apoyo se entiende de un modo asimétrico. Lleva la asimetría al interior de la teoría de la justificación del modo preciso que al coherentista le gustaría evitar. (Ideas similares, por ejemplo la de que nuestras creencias sensoriales son nuestras evidencias o nuestros datos, tienen el mismo efecto.) El requisito de que lo sensorial preste apoyo a lo no sensorial equivale a decir que la justificación es unidireccional, yendo de lo sensorial a lo no sensorial, y, por ello, a decir que la justificación adopta dos formas: en primer lugar, la justificación de lo no sensorial por lo sensorial y, en segundo lugar, la (de algún modo diferente) justificación de lo sensorial. Lo que equivale a abandonar la tesis monista esencial del coherentismo en favor de alguna forma de fundamentalismo, por muy limitada que esa forma resulte ser.

El coherentismo podría, por supuesto, tratar de escapar a este ataque pretendiendo que la mera distinción entre las creencias sensoriales y las no sensoriales no equivale a imponerle una de las asimetrías no deseadas. Pero creo que, por ese camino, no tiene escapatoria. En primer lugar, no es la distinción misma la que crea asimetría, sino la

demanda de que, distinguidas de ese modo, las creencias sensoriales soporten a las no sensoriales. En segundo lugar, parece que hay razones poderosas para que incluso el coherentista adjudique a las creencias sensoriales algún papel especial en la epistemología del individuo. No todas esas razones deben tener el mismo peso, pero mencionaré tres de ellas.

En primer lugar, aquellos objetos cuya justificación estamos considerando son conjuntos de creencias, y todos los conjuntos de creencias con los que estamos familiarizados (el nuestro propio y el de nuestros contemporáneos) están, de hecho, basados en la experiencia. No tenemos razón alguna para preocuparnos, por tanto, con inexistentes conjuntos de creencias que carecen de fundamento empírico. En segundo lugar, para que un conjunto cuente como conjunto de *creencias* —para que la actitud proposicional implicada sea la creencia— debe ser concebido como una respuesta a los estímulos del medio. No se trata de que las creencias completamente desconectadas de la experiencia estarían injustificadas; se trata más bien de que, en ese caso, no serían creencias en absoluto. En tercer lugar, parece posible, aunque la cuestión deba determinarse empíricamente, que la confianza asimétrica en la experiencia que se expresa en la idea de que nuestra evidencia es la experiencia produciría, a partir del mismo *input*, conjuntos de creencias con una mayor coherencia, y, si esto es así, hay razones para introducir una asimetría en la explicación desde el mismo proyecto coherentista.

Dado que debe existir alguna asimetría, ¿puede vérselas con ella el coherentista? Podría intentarlo distinguiendo entre dos tipos de seguridad en las creencias, la antecedente y la subsiguiente. La *seguridad antecedente* es aquella que una creencia lleva consigo, que es anterior a cualquier consideración sobre lo bien que la creencia pueda ajustar con otras creencias o sobre la coherencia del conjunto. Podríamos mantener que las creencias sensoriales tienen un grado de seguridad antecedente al ser *prima facie* fiables o justificadas; habrá, hasta la infalibilidad, grados mayores de seguridad antecedente. La *seguridad subsiguiente* es la seguridad que adquiere una creencia como resultado de su contribución a la coherencia del conjunto. Todas las creencias justificadas, en una explicación coherente, tienen cierto grado de seguridad subsiguiente.

Los coherentistas tradicionales pretenderían que ninguna creencia tiene una seguridad antecedente mayor que otra. Podríamos denominar esta posición puro coherentismo; una forma extrema de ella mantiene que ninguna creencia tiene ninguna seguridad antecedente en

absoluto. Pero el argumento anterior podría convencernos de que las creencias sensoriales tienen una seguridad antecedente de la que carecen otras creencias. Si este «coherentismo débil» es consistente, es posible que pueda afrontar las demandas del empirismo. Pero parece que el coherentismo débil corre el peligro de ser sólo otro rótulo para una forma de fundamentalismo. Después de todo, la fiabilidad *prima facie* y características semejantes se mencionaron en 4.3 como cruciales de las formas no-clásicas de fundamentalismo. Por tanto, ¿es posible ser empirista y aceptar una relación asimétrica entre las creencias sensoriales y las otras, sin ser por ello un fundamentalista? En la precedente discusión sobre Quine (7.2 y 3) parecía que la insistencia empirista en que nuestras creencias se fundamentaran en la experiencia debería ser, de algún modo, compatible con el holismo completo tanto en epistemología como en teoría del significado. ¿Podemos mostrar con más detalle cómo sería posible tal cosa?

8.5. COHERENTISMO Y EMPIRISMO

El punto de vista coherentista más fructífero puede encontrarse en la obra de F. H. Bradley. Bradley es un empirista, que se expresa a este respecto con toda la claridad que pudiera desear el empirista más fervoroso:

> Estoy de acuerdo en que dependemos vitalmente del mundo sensorial, en que nuestros materiales provienen de él, y en que sin él no podría comenzar el conocimiento. Estoy de acuerdo en que siempre tenemos que volver a este mundo, no sólo para acceder a cosas nuevas, sino para confirmar e incrementar las viejas [Bradley (1914), p. 209].

Vemos aquí que Bradley adjudica a los «datos de la percepción» o al «mundo sensorial» un papel asimétrico en la epistemología del individuo. De hecho, se trata de una asimetría compleja. En parte es genética, el material *proviene del* mundo sensorial, y sin ese mundo no podría comenzar el conocimiento. También tiene un papel continuo, tanto por nuestra necesidad de volver constantemente a los «datos de la percepción» previos, como por nuestra necesidad da dotar de sentido al flujo constante de nueva vida sensorial. Esta asimetría compleja refleja (si fuera posible alterar el orden temporal) los argumentos de Quine en favor de la teoría verificacionista del significado; éstos eran o bien sobre requesitos genéticos, como cuando escribe sobre el tipo de significado que es básico para el aprendizaje del propio lenguaje, o

bien sobre requisitos continuos, como cuando escribe sobre el tipo de significado básico para la traducción (7.2).

Bradley aceptaría que el mundo sensorial desempeña un papel especial en epistemología, pero no que ese papel equivalga al tipo de asimetría que caracteriza el fundamentalismo (ibíd, p. 210).

> Para comenzar mi construcción, considero como absoluto el fundamento [...]. Pero no se sigue de ello que mi construcción descanse continuamente en los comienzos de mi conocimiento. Puesto que es otro el sentido en el que mi mundo descansa sobre los datos de la percepción.

Bradley mantiene que la experiencia proporciona datos (asimetría genética) pero que la cuestión de si algo que parece un dato debe permanecer como un hecho aceptado es una que no está determinada, ni siquiera parcialmente, por su origen como dato. La aceptación de los datos en nuestro mundo se produce del mismo modo y por los mismos criterios que la de cualquier otra proposición. En cada caso, la prueba crucial es lo que él denomina «sistema», o, en otras palabras, el que la coherencia de nuestro mundo se incremente o no por su admisión como un hecho. En este respecto, no hay ninguna asimetría; todas las proposiciones (en el sentido, por decirlo de algún modo, de propuestas) que están justificadas reciben exactamente el mismo tipo de justificación.

¿Es consistente la posición de Bradley al aceptar una asimetría mientras que rechaza otra? Podría argumentarse en su contra que, incluso si estamos de acuerdo en que tanto las proposiciones que tratan de datos como las demás se justifican por su contribución al sistema, todavía queda una asimetría crucial que no es genética. Recordemos que el primer objetivo del sistema era la necesidad de dotar de sentido al mundo sensorial; incluso aunque en la ejecución de esa finalidad rechazáramos algunos de los elementos de ese mundo, todavía existiría una asimetría en el propósito de la sistematización. La asimetría se revela en la exigencia de aceptar por lo general los elementos que sean datos. Sería una objeción contra cualquier sistema el que requiriera un rechazo sustancial de los «datos de la percepción», independientemente de si con ello se incrementa la coherencia del sistema.

Por supuesto, sería posible la escapatoria fácil de argumentar que la objeción es sólo válida contra el coherentismo puro, que sostiene que todas las creencias tienen la misma seguridad antecedente; y que no afecta al coherentismo débil, que acepta que algunas creencias tienen mayor seguridad antecedente que otras y puede, por tanto, ofrecer un análisis de la necesidad de que, en general, los elementos conside-

rados como datos hayan de sobrevivir al escrutinio epistemológico. Pero no sería ésta una estrategia adecuada, dado que la cuestión relevante es otra. Lo realmente importante es si ese tipo de seguridad antecedente, si estamos obligados a aceptarla, equivale o no a una asimetría en el análisis que demos de la justificación y, por tanto, a una teoría bipolar de la justificación del tipo que sólo puede proporcionar un fundamentalista. Si es equivalente, tenemos un argumento tan efectivo contra el coherentismo débil como contra el puro. De modo que hay dos cuestiones distintas aquí: ¿es necesaria alguna forma de seguridad antecedente para las creencias sensoriales?, y, si eso es así, ¿introduce la seguridad antecedente una asimetría que nos obligue a admitir una teoría bipolar de la justificación?

La seguridad antecedente de la que gozan las creencias sensoriales parece equivaler al hecho de que debemos aceptarlas como verdaderas si nada cuenta en su contra. Pero ¿no hacemos lo mismo, sensatamente, con todo aquello que aceptamos como creencia? Nos quedamos con cualquier creencia a menos que haya alguna razón para rechazarla. De modo que, en este sentido, *todas* las creencias tienen seguridad antecedente. Y ello no introduce dos formas de justificación. No hay ninguna asimetría originada en la aceptación de que todas las creencias tienen algún grado de seguridad antecedente, siempre que la seguridad antecedente de la que gocen sea del mismo tipo en todos los casos.

Pero es posible que el problema sea que creencias diferentes tienen grados diferentes de seguridad antecedente, y que es típico del empirismo el mantener que las creencias sensoriales tienen un grado *mayor* que otras. ¿Puede el coherentista dotar de sentido a esta idea en sus propios términos? El problema parece ser que el hecho de que una creencia pueda ser así más segura que otra es un hecho previo e independiente de cualquier consideración sobre la coherencia con otras creencias, introduciendo de nuevo una asimetría para la que no puede existir una explicación coherentista.

El problema, por tanto, es si el coherentista *puede* ser empirista, no el de si *debe* serlo. Y el empirista se caracteriza por la actitud que adopta hacia sus creencias sensoriales; exige más de lo que exigiría otro para estar dispuesto a rechazarlas. Pero, si esta actitud es extrínseca a las creencias sensoriales mismas y si puede ser considerada legítimamente como una creencia adicional, se trata de una creencia que podría compartir el coherentista. Y, si la comparte de hecho, se obtendrá el resultado apetecido. La eliminación de la creencia sensorial crearía una distorsión mayor y requeriría más justificación, por el mero

hecho de que la creencia típicamente empirista es también parte del conjunto de creencias. Por ello, las creencias sensoriales del coherentista tendrán un grado mayor de seguridad, pero será una seguridad subsiguiente, no antecedente; dado que, normalmente, será considerada de un modo compatible con el coherentismo, es decir, en términos de la coherencia interna del conjunto de creencias.

Si esto es correcto, el coherentismo puro es más fuerte que el débil. Si el coherentismo débil ha de caracterizarse por su disposición a admitir grados diferentes de seguridad antecedente, su posición es genuina e innecesariamente débil.

Por todo ello, debemos concluir que el coherentismo es compatible con el empirismo. Una cuestión diferente es la de si el coherentista *debe* ser empirista. Se hablará de ello en el capítulo 11. Pero el coherentista parece tener perspectivas halagüeñas en este punto. Para él, es una cuestión empírica el que, al final de la jornada, la adopción de la actitud empirista hacia las creencias sensoriales dé como resultado un conjunto más coherente. Es una cuestión empírica el que esta especie de tozudez empirista produzca o no dividendos en último término. Y ésta es la manera en la que el coherentista *debe* tratar de justificar el empirismo.

LECTURAS COMPLEMENTARIAS

«On truth and coherence», cap. 7 de Bradley (1914).

Rescher (1973, cap. 2) distingue entre teorías «criteriales» y «definicionales» de la verdad, y rechaza la pretensión de la teoría de la coherencia de ser definicional. No está claro que una teoría «criterial» de la verdad sea reconociblemente distinta a una teoria definicional de la justificación.

Firth (1964) discute la cuestión de si una explicación coherentista del conocimiento puede dotar de sentido a la idea de que el conocimiento empírico se basa en la experiencia; la defensa que ofrece parece ser una forma de coherentismo débil.

Seilars (1979) responde a Firth de un modo más consistente con el puro coherentismo.

Ewing (1934, cap. 5) es un critico de las teorías de la coherencia que simpatiza con ellas. Está de acuerdo con Rescher en que la teoría de la coherencia de la verdad debe considerarse criterial más que definicional.

Blanshard (1939, caps. 25-27) ofrece la formulación autorizada más reciente de una posición descaradamente coherentista.

La idea de lo coherente como lo mutuamente explicativo parece que se deriva de Sellars(1936, pp. 321-358).

Una extensión del argumento de 8.4-5 se encuentra en Dancy (1984a).

Cornman (1977) sugiere que una mezcla de fundamentalismo y coherentismo resultará ser «la teoría de la justificación empírica más razonable».

9. COHERENCIA, JUSTIFICACIÓN Y CONOCIMIENTO

9.1. EL ARGUMENTO DEL REGRESO DE JUSTIFICACIONES

¿Qué tiene que decir el coherentista con relación al argumento del regreso de justificaciones (4.1)? El argumento, después de todo, se supone que muestra que la distinción entre justificación inferencial y no inferencial no es opcional; estamos obligados a ella, so pena de caer en un regreso al infinito y en el escepticismo. La dificultad en este punto radica en que, como lo expresamos anteriormente, el argumento del regreso dependía de una concepción de la justificación que era a la vez no-holista y lineal. (Era lineal en el sentido de suponer que A podría estar justificada por B, B por C, etc., o bien *ad infinitum* o bien cayendo en circularidad.) Por tanto, es preciso reformular el argumento en términos de la concepción holista de la justificación que es central al coherentismo. ¿Es posible tal cosa? De hecho, soy incapaz de encontrar ninguna reformulación del argumento del regreso de justificaciones que lo haga efectivo en contra del coherentismo.

El análisis coherentista de la justificación puede ser formulado del siguiente modo: una creencia B1 está justificada en la medida en que contribuya a la coherencia del conjunto de creencias del que es un miembro. Supongamos que el conjunto contiene como miembros las creencias B1 – B*n*, y supongamos también que el tipo de justificación bolista en el que estamos interesados puede ser considerado razonablemente inferencial. Podría ser incluso inductivo, en la medida en que consideráramos la inducción como una inferencia a la mejor explicación [como se argumenta en Harman (1970)]. ¿Nos enfrentamos ahora al mismo peligro de que, a menos que exista otra forma de justificación, toda justificación será condicional (4.1)? Parece que hay dos formas en que podría plantearse esta dificultad.

La primera es que B1 sólo esté justificada si lo está (por ejemplo) B2, etc. Pero tal cosa no se sigue de la concepción coherentista. Incluso si B2 no estuviera justificada, y existiera una alternativa B2' que tuvie-

ra una contribución mayor a la coherencia del conjunto B1 – B*n*, todavía podría suceder que B1 estuviera justificada. Lo estaría si, en líneas generales, ninguna alternativa hiciera el conjunto más coherente. Parece, pues, que no hay ningún sentido en el que la justificación de B1, incluso aunque exista por apelación a otros miembros del conjunto y al papel que desempeñan en la construcción del mismo, esté condicionada a la *justificación* de otros miembros. Está, por supuesto, condicionada a su presencia, pero este tipo de condicionalidad no crea ningún regreso al infinito que sea vicioso.

La segunda alternativa sería la de sostener que B1 debe contar como creencia justificada por apelación a su contribución a la coherencia del todo. Pero su papel en el todo sólo podría justificar B1 si el mismo todo estuviera justificado. La justificación de B1 sería de nuevo condicional, pero ahora lo sería respecto a la justificación del conjunto del que es miembro. Creo que hay dos respuestas distintas que, en último término, coinciden. La primera sería decir que este tipo de condicionalidad no importa, dado que confiamos en que el conjunto sea incondicionalmente (realmente) coherente y, por ello, en que esté realmente justificado. No habrá, por tanto, peligro alguno de una cadena al infinito, dado que en el primer eslabón ya no nos encontramos con algo cuya justificación no está condicionada. Sin embargo, esta contestación no me deja completamente satisfecho, puesto que podría mantenerse que admite la existencia de dos formas distintas de justificación, la de las partes y la del todo. Una respuesta menos problemática es la de decir que el coherentismo no otorga sentido alguno a la idea de la justificación de un conjunto completo. Un miembro está justificado por su contribución a la coherencia del conjunto; la justificación se define en términos de coherencia, pero la justificación y la coherencia no son lo mismo (cf. las observaciones al final de 8.1).

Ésta no es una cuestión meramente verbal, un exabrupto coherentista o un rechazo injustificado del rótulo de «justificación» para algunos casos. Tras ello yace una explicación más sustantiva de la práctica de buscar justificaciones y de nuestro modo de actuar en lo que Levi denomina «la empresa del conocimiento» [Levi (1980)]. En cualquier momento nos podemos encontrar con tensiones, si no contradicciones, en el seno de un enorme conjunto de creencias. La cuestión importante es la de qué hacer entonces. Lo que buscaremos es un reajuste que incremente la coherencia global. Cualquier creencia puede ser dejada de lado o reemplazada, pero todas las creencias están codificadas en otras, de modo que su rechazo creará problemas adicionales. Una creencia se acepta como justificada si no hay ninguna otra aparente alter-

nativa que pudiera comportarse mejor. Pero no hay sentido alguno en el que se evalúe la totalidad del conjunto de creencias, en función de su justificación, para su aceptación o su reemplazamiento. Lo mejor, de hecho lo único, que podemos hacer es mejorar la coherencia interna, usando la amplia estructura del conjunto para evaluar la contribución de los miembros individuales y de los nuevos candidatos. Si es esto lo que constituye el proceso de justificación, difícilmente puede sorprendernos que no tenga ningún sentido la cuestión de si el conjunto entero está o no justificado.

Parece, pues, que los coherentistas pueden escapar a la conclusión de que las creencias justificadas «inferencialmente» sólo están justificadas de un modo condicional. Pueden mantener que cada creencia está (o debe estar) justificada de hecho y de un modo no condicional gracias a su contribución a la coherencia del conjunto. Si esto es así, estamos obligados a considerar el argumento del regreso de justificaciones no como un argumento independiente en favor del fundamentalismo, sino como expresión de la actitud fundamentalista sobre la justificación, una expresión que pone de manifiesto las consecuencias de considerar que la justificación es no–holista y lineal.

No obstante, todavía nos aguarda una regresión diferente. He sugerido que una creencia puede estar justificada por el hecho de que hace determinada contribución al conjunto de creencias del que es miembro. Pero ¿puede bastarme realmente como justificación para mantener una creencia el que incremente la coherencia de mi conjunto de creencias? ¿Qué sucede si no tengo ni idea de que ése es el caso? Ciertamente, sólo estaría justificado para mantener esa creencia si, de algún modo, creo que contribuye del modo adecuado. E incluso esto puede no ser suficiente. Podría darse el caso de que no estuviera justificado para creer que la contribución es la adecuada; esta segunda creencia podría resultar completamente gratuita o podría mantenerse por motivos completamente irracionales. ¿No es necesario también que yo esté justificado al creer que la primera creencia contribuye a la coherencia del conjunto?

Quizá pueda verse más fácilmente esta regresión desde un punto de vista formal. Si q = «la creencia de a de que p contribuye a la coherencia del conjunto de creencias de a», parece que estamos en peligro de definir CJap en términos de:

1. q
2. Caq
3. CJaq.

Con lo que generamos un nuevo regreso de justificaciones. Para cada creencia justificada *p* debe existir una creencia justificada adicional *q*, y así *ad infinitum*. En el ejemplo anterior, considerando *r* como «la creencia de *a* de que *q* contribuye a la coherencia del conjunto de creencias de *a*», estamos obligados a continuar

4. *r*
5. C*ar*
6. CJ*ar*,

y, una vez efectuado este movimiento, lo tendremos que repetir indefinidamente. Este nuevo regreso al infinito deriva de un punto de vista internalista sobre la justificación. Nos debe recordar el que mencionamos en 3.5. Para ver si se trata de una amenaza al coherentismo, necesitamos considerar con más detalle tanto el contraste entre internalismo y externalismo como las supuestas ventajas del externalismo.

9.2. INTERNALISMO Y EXTERNALISMO

El regreso internalista de justificaciones genera un problema similar para el coherentista al que no afectaba el anterior argumento del regreso. Pero para el fundamentalista las cosas son mucho peores. Se diría que el regreso internalista tendrá el efecto de que el regreso anterior no puede detenerse, por muchas sutilezas que inventemos. Supongamos que alcanzamos algunas creencias básicas, y supongamos que lo que las hace capaces de estar justificadas de un modo no inferencial es cierta propiedad epistémica E (que podría ser cualquier cosa desde la mera infalibilidad hacia abajo). El internalismo nos empuja a sostener que la creencia básica que tiene *a* de que *p* sólo puede justificarse por apelación a E si a cree que su creencia posee la propiedad E; y nos empuja, además, a admitir que también es necesario que *a* esté justificado en su creencia de que su creencia posee E. Si se concede todo esto, ya no hay manera de detener el regreso de justificaciones. Ni siquiera podría detenerlo el descubrimiento de algunas creencias justificadas de un modo no inferencial, dado que sólo pueden justificarse en presencia de una creencia justificada ulterior. Por lo que ningún tipo de creencia pasará nunca de ser una creencia justificada condicionalmente.

Todo esto nos proporciona un buen incentivo para iniciar la retirada, aceptando el externalismo y reduciendo los requisitos para la justi-

ficación. El externalista lo tiene más fácil de dos maneras distintas. En primer lugar, no tiene problema alguno con el argumento habitual del regreso de justificaciones. Consideremos, como ejemplo de una posición externalista, la siguiente definición de CJ*ap:*

CJ*ap* ≡ la creencia de *a* de que *p* se adquirió por un método fiable.

Dado que la inferencia a partir de la creencia justificada es un método fiable de adquisición de creencia, tanto las creencias justificadas de un modo inferencial como las justificadas no inferencialmente pueden ser subsumidas bajo esa definición. La justificación inferencial no es el territorio particular del internalismo. En este punto, le será fácil al externalismo encontrar un análisis satisfactorio de la creencia externamente justificada. No hay necesidad de retrotraerse al nivel básico, como se pensó tradicionalmente. Si suponemos, como es verosímil suponer (al menos en ciertos casos), que preguntar a los padres es un método fiable de adquisición de creencias, cualquier creencia obtenida de esta manera contará como creencia justificada de un modo no inferencial. El argumento del regreso de justificaciones ya ha perdido aquí mucha de su mordiente, aunque todavía pueda llevarnos a la distinción entre justificación inferencial y no inferencial.

En segundo lugar, el externalismo parece no tener problema alguno con nuestro segundo argumento escéptico, el del error (1.2). El mayor atractivo de este argumento era que debo señalar una diferencia entre el caso actual y otros casos relevantemente similares en los que estaba equivocado antes de poder decir que ahora sé. Pero esta exigencia parece estar basada únicamente en una perspectiva internalista. El externalista se siente feliz con suponer que, en la medida en que de hecho haya una diferencia entre este caso y otros, yo puedo saber ahora incluso aunque sea incapaz de señalar esa diferencia. Si esto es correcto, todo lo que consigue el segundo de nuestros argumentos escépticos es expresar una dificultad general añadida para los análisis internalistas del conocimiento y la justificación.

Dadas esas ventajas del externalismo, ¿qué motivos puede tener alguien para ser internalista? La respuesta es que los filósofos son internalistas sólo porque creen que deben serlo. Antes de que podamos sentir confianza en la opción externalista, debemos estar bien seguros de haber entendido bien los supuestos del internalismo. Y, antes de hacer tal cosa, es preciso que distingamos diversos grados de internalismo que aparecen de un modo más bien confuso en la literatura contemporánea.

9.3. GRADOS DE INTERNALISMO

Supongamos que existe, por lo menos, una cláusula *c* en la definición de S*ap* o de CJ*ap* que no comienza con C*a*– o con S*a*–. Antes ya se dio un ejemplo: la creencia de *a* de que/? se adquirió por un método fiable. El internalismo podría estar constituido por cualquiera de estas afirmaciones:

1. no debe haber tal tipo de cláusula *c*;
2. para cada cláusula *c* debe existir una cláusula S*ac*;
3. para cada cláusula *c* debe existir una cláusula C*ac*;
4. para cada cláusula *c* debe existir una cláusula CJ*ac*;

Hay diferentes motivos para mantener o negar cada una de ellas. 1 y 2 son las más extremas. Con mucho, la menos exigente es 3, y argumentaré que debemos preferirla, o, al menos, que debemos preferirla a 4. Ofreceré, a continuación, algunos comentarios sobre cada una de ellas.

1. Literalmente, eliminaría cualquier definición contemporánea de conocimiento, porque en todas ellas la primera cláusula viola esta condición. Podríamos restringir el alcance de 1 al análisis de la justificación; pero tal cosa la inutilizaría, como lo harían otros muchos compromisos. Entonces, ¿de qué sirve 1? Quizá la justificación, la evaluación de las creencias para su aceptación o su rechazo, es una actividad que debe ser ejecutada por el individuo. Pero el individuo no puede distinguir en cada caso particular entre sus creencias y la verdad. De modo que, cuando hay una cláusula C*ap* y una cláusula separada *p*, una al menos es redundante. Si esto yace tras la aseveración de 1, sería razonable restringirla al análisis de la justificación sobre la base de que el individuo no está en una posición privilegiada a la hora de determinar lo que sabe. Hasta esto iría en contra de una firme tradición filosófica, que comienza en Platón y reaparece de un modo llamativo con Descartes. Si, a pesar de todo, la aceptamos momentáneamente, quizá podamos distinguir otros motivos impulsores tras 1, incluso en esta versión restringida, que provienen de la filosofía de la mente. He aquí dos opiniones sobre lo que determina la naturaleza de un estado mental como la creencia. La primera opinión (la análoga del internalismo) mantiene que la naturaleza de la creencia está completamente determinada por sus características subjetivas, por lo que el creyente goza de privilegios exclusivos para poder decirnos la naturaleza

de sus creencias. El punto de vista contrario (análogo al externalismo) es el de que la naturaleza de ciertas creencias está determinada por algo distinto a sus características subjetivas; a veces puede ser crucial la naturaleza del mundo circundante. No se trata de la consideración débil de que la naturaleza del mundo causa la naturaleza de la creencia, sino de la consideración fuerte de que la naturaleza de la creencia depende lógicamente de la naturaleza o la existencia de algún rasgo del mundo. Un ejemplo será útil. ¿Qué es lo que hace que una creencia lo sea sobre una olla particular? ¿Es el «ser sobre» una característica subjetiva interna de la creencia, una característica que la creencia podría llevar consigo con total independencia de la naturaleza o la existencia de cualquier objeto (la olla, quizá)? O, ¿depende lógicamente este «ser sobre», a pesar de ser una característica esencial de la creencia, de la existencia o la naturaleza del objeto apropiado? En el caso de ciertas creencias sobre objetos particulares (*de re*), no hay ningún residuo interno y separado que pudiéramos considerar como la parte «interna» de la historia y suponer que continúa presente incluso cuando no existiera el objeto relevante. No hay ningún factor común de nivel más bajo compartido por las creencias *de re* que tienen éxito y las que no lo tienen.

Este tipo de externalismo en filosofía de la mente es discutible, pero es probable que tenga consecuencias interesantes en teoría de la percepción (que se discutirán, posteriormente, en 11.4). Si lo adoptamos, ¿dónde nos llevará en teoría del conocimiento y de la justificación? Creo que no muy lejos. Lo que sí hace es alejarnos del deseo de pensar que todo lo relevante para la justificación debe ser transparentemente accesible al que crea; puesto que, si lo que cree depende no sólo de sus estados subjetivos sino del estado independiente del mundo, no tiene sentido continuar insistiendo en que todas las cláusulas en la definición de justificación deban comenzar con S*a*– o C*a*–. Podríamos haber esperado con ello asegurar que la justificación fuera una actividad del individuo, pero ésta era una vana esperanza. Dado que el procedimiento no puede alcanzar lo que intentábamos, daría lo mismo no haberlo iniciado.

Este tipo de externalismo sólo es discutible en teoría de la justificación. En teoría del conocimiento parece ser verdadero, aunque bajo ciertas lecturas pueda llevar al rechazo de una de las condiciones del conocimiento: la de la creencia. El conocimiento no se consideraría entonces como creencia con algo añadido, sino como un estado mental independiente. Pero esto nos llevaría demasiado lejos. Por el momento, parece que, si hay un debate entre las concepciones inter-

nalista y externalista del conocimiento, todavía no lo hemos localizado.

2. Esta versión del internalismo proviene de Armstrong (1973, pp. 153-157), aunque está anticipada en el *Teeteto* de Platón (209E-210B). Sin embargo, tal y como la presentamos, sólo se refiere a la teoría del conocimiento. No hay ninguna sugerencia de que para estar justificado sea preciso *saber*, por ejemplo, que el método seguido es fiable. Más aún, tampoco resulta convincente en teoría del conocimiento. También crea un regreso al infinito. Sugiero que nos olvidemos de ella.

3. Ésta es la teoría menos exigente. No puede ser atacada sobre la base del externalismo en filosofía de la mente, dado que ese tipo de externalismo sólo argumentaba que no tenía sentido insistir en que todas las cláusulas comiencen con C*a*– o S*a*–. Y no crea ningún regreso al infinito, a menos que se conjunte con 4. (Tales regresos en 3.5 y 9.1 eran producto de la combinación de 3 y de 4.) Parece también aplicable al conocimiento y a la justificación, y dota de sentido al internalismo, dado que requiere que el creyente «internalice» los hechos que hacen que su creencia esté justificada, o que su creencia verdadera sea conocimiento. Si hay un sentido de justificación «interna» en el que todo lo que podemos pedir a alguien es que retenga las creencias que, según él, satisfagan las condiciones para la justificación, tenemos una forma de internalismo que nos ofrece una distinción relevante. La justificación interna se da cuando todas las cláusulas relevantes C*a*– son verdaderas; la justificación externa se da cuando todas las otras cláusulas son también verdaderas.

Podría parecer que esta teoría es tanto externalista como internalista, dado que da sentido a una idea de justificación externa, *de facto*, de una creencia que es diferente de la idea de justificación interna. Pero esto sería un error, por dos razones. En primer lugar, de acuerdo con la teoría, la justificación externa incluye la interna. Podemos tener la última sin la primera, pero no al revés; de modo que no podemos dotar de sentido a la idea de que alguien pueda estar justificado en una creencia aunque no sea consciente de algunos de los hechos en virtud de los cuales está justificado. En segundo lugar, esta teoría internalista se opone claramente a una teoría externalista como la de Nozick. Hay en este punto una especie de transición gradual, y esta teoría se halla en algún lugar del medio.

Una de las ideas que subyace a esta teoría es una analogía entre

epistemología y ética. Cuando consideramos nuestra actitud hacia la acción moral, ¿nos basta con que la acción sea de hecho correcta, o exigimos algo más? Ciertamente, también exigimos que el agente crea que la acción es correcta. En caso contrario, la acción no tiene propiedades morales en absoluto; es tan sólo afortunada, más que merecedora de aprobación moral. De un modo similar, el internalista en epistemología podría exigir que la justificación externa comience con, e incluya, la interna.

Considero, pues, que esta teoría más débil es una forma reconocible de internalismo y satisface, al menos, algunas de las ideas a partir de las que surgen las diferentes formas de internalismo.

4. Esta teoría más exigente es la que genera el regreso al infinito. Podemos proseguir con la analogía en ética para mostrar por qué, sin embargo, parece atractiva e incluso imprescindible. Ya hemos sugerido que una acción no puede ser correcta a menos que su agente crea que es correcta. Sin embargo, ¿no debemos añadir algo más? ¿Qué sucede si la creencia del agente es pura coincidencia? Podría suceder que sus razones para pensar que es correcta, si es que las tiene, fueran absolutamente irrelevantes; por ejemplo, podría pensar que es correcta porque se hace en martes. Por tanto, podría parecer que es necesario añadir el requisito de que la creencia del agente respecto a la corrección de la acción ha de estar justificada. Con lo que nos vemos conducidos a la forma presente y más extrema del internalismo. En epistemología nos encontramos con las mismas ideas. Si exigimos del agente que crea que su primera creencia se adquirió, por ejemplo, por métodos fiables, con toda certeza deberemos exigir que esta creencia no sea una mera coincidencia, que tenga buenas razones y, por tanto, que la creencia esté justificada.

Aunque esta línea de pensamiento tiene cierta fuerza, está equivocada. Lo que de hecho nos lleva desde la teoría 3 a la 4, como al final del parágrafo anterior, es la idea de que la justificación y el conocimiento deben depender, de algún modo, de algo que no sea suerte o coincidencia. Esto era lo esencial en los contraejemplos de Gettier (2.2); en la definición tripartita del conocimiento no había nada que excluyera el conocimiento por suerte. La insistencia de que una cláusula C*a*J*a*- correspondiente a cada cláusula C*a*– es un intento de resolver el problema de Gettier y de eliminar el conocimiento por mera suerte. La insistencia en que una cláusula C*a*CJ*a*- corresponda a cada cláusula *Cía*- es nuestra teoría internalista número 3. Si las unimos, nos topamos con el regreso vicioso al infinito, como vimos en 3.5.

La solución, por tanto, debe ser la de encontrar otra respuesta al problema del conocimiento por mera suerte; si lo logramos, no tenemos ninguna necesidad de añadir la teoría 4 a la 3. El regreso vicioso al infinito no es una consecuencia del internalismo sólo, sino del internalismo añadido a algo relativamente distinto.

Nuestra conclusión hasta ahora es que hay una forma de internalismo que no es vulnerable al regreso internalista de 9.1. Aunque el externalista ofrezca ventajas contra el escéptico, el internalismo sigue siendo una opción abierta.

9.4. INTERNALISMO Y COHERENTISMO

¿Tienen los coherentistas la posibilidad de elegir entre internalismo y externalismo? Parece que sí. Las relaciones entre los miembros del conjunto de creencias de uno mismo que hacen coherente el conjunto pueden existir con independencia de que uno las crea o no. De modo que, si podemos aceptar que el internalismo no es necesario, el coherentista tiene acceso a las ventajas del externalismo.

Hay una pequeña duda respecto a esta solución tan reconfortante, dado que, parece obvio, las condiciones requeridas por las formas extremas de internalismo tenderían, si se tuvieran presentes, a incrementar la coherencia de un conjunto de creencias dado. Los coherentistas piensan en términos de un incremento de justificación a medida que el conjunto de creencias se hace coherente. Y parece que la coherencia del conjunto se incrementará siempre por la adición de la creencia de que existen relaciones relevantes de explicatividad mutua. Es decir, si tales relaciones existen, uno siempre estará justificado para creer que si existen. Sin embargo, ello sólo muestra que la justificación se incrementa por el desplazamiento desde la perspectiva externalista a la internalista, y no que nunca esté presente sin la adición internalista. Hasta los externalistas más extremados pueden estar de acuerdo en que hay cierta ganancia en la internalización de los hechos relevantes; y el coherentista se encuentra en una posición bastante buena para mostrar que esto es así.

Si nos encontramos ante una opción abierta, ¿qué debemos elegir? Los intentos más importantes de argumentar en favor de una u otra de estas posiciones [Goldman (1980) en favor del externalismo, y Bonjour (1980) en favor del internalismo] captan de una forma bastante insegura el contraste entre las dos posiciones y muestran la tendencia a sustentarse en la analogía entre ética y epistemología que ya

hemos considerado al distinguir entre diferentes formas de internalismo [cf. también Pollock (1979), pp. 108-111]. Pero es improbable, por dos razones diferentes, que el uso de esta analogía resulte efectivo. En primer lugar, es bastante fácil poner en cuestión la fuerza de la analogía. En segundo lugar, la relación entre la corrección «objetiva» y la «subjetiva» en ética (análoga a la justificación externa e interna) no ha sido todavía comprendida bien y se encuentra en el centro de una acalorada disputa.

Así pues, parece que hay que aceptar que no hay argumentos decisivos en ninguno de los dos bandos. Mi propia opinión es la de que el internalismo se encuentra mejor respaldado por nuestras intuiciones. Lo que quiere decir que no puedo quedarme satisfecho con la dependencia exclusiva en la posibilidad del externalismo como respuesta al argumento escéptico del error; esa posición será intrínsecamente inestable. Todavía necesitamos una estrategia que pueda utilizar el internalista y que eluda el argumento del error.

9.5. COHERENTISMO, REALISMO Y ESCEPTICISMO

¿Nos proporciona el mismo coherentismo este tipo de estrategia? El argumento del error se expresó en términos realistas; como el primero de los argumentos escépticos, parecía depender de diferencias reales e irreconocibles (trascendentes a toda evidencia). Si tal cosa es así, y el coherentismo resultara ser incompatible con el tipo de realismo que se necesita, un coherentista tendría una perspectiva desde la que el argumento del error ya no constituiría ninguna amenaza.

De modo que la cuestión es la de si el coherentismo es una forma de antirrealismo. Para responderla, es preciso que proporcionemos un mapa filosófico de las diferentes formas de antirrealismo. Hay cuatro grados de antirrealismo, que van desde el solipsismo extremo, a través del idealismo y del fenomenalismo, a lo que denomino antirrealismo puro.

El solipsista (como se le define en este punto) sostiene que todas las proposiciones imaginables tratan de sus propias experiencias. El solipsista más estricto sólo podrá imaginar sus propias experiencias actuales; solipsistas más laxos aceptarán tanto las actuales como las pasadas, las actuales y las futuras o las pasadas, presentes y futuras.

El idealista sostiene que todas las proposiciones imaginables tratan de sus propias experiencias o de las de otros; y, de nuevo, existen variantes relativas al tiempo: se puede escoger el presente, el presente y pasado, presente, pasado y futuro, etc.

El fenomenalista sostiene que todas las proposiciones imaginables tratan o bien de experiencias reales de sí mismo o de los otros, o bien de experiencias que él o los demás podrían tener en determinadas circunstancias. El fenomenalista habla de experiencias reales o posibles, mientras que el idealista se limita a las reales. De nuevo, hay varias versiones de fuerza diferente de acuerdo con las diferencias temporales entre las experiencias, reales o posibles.

El antirrealista puro sostiene que algunas proposiciones imaginables tratan de asuntos distintos a las experiencias, reales o posibles, de él mismo o de los demás, pero que, sin embargo, no entendemos que tales proposiciones puedan ser verdaderas de un modo *irreconocible*. De modo que el antirrealista puro, como otros antirrealistas, niega la posibilidad de una verdad que trascienda a toda evidencia posible, aunque admite que algunas verdades conciernen a cosas como los objetos materiales, que no pueden suponerse reductibles a la experiencia. Un antirrealista, pues, puede admitir la existencia del mundo material, siempre y cuando insista en que carece de propiedades trascendentes a toda evidencia.

También hay distintos grados de antirrealismo puro, dependiendo de lo que se entienda por «irreconocible»; debemos preguntarnos: «irreconocible ¿para quién y cuándo?». Las respuestas pueden variar: la más estricta sería: «irreconocible para mí, ahora». La más débil: «irreconocible para un ser con capacidades cognitivas mucho mayores que las nuestras, como lo sería Dios».

Del mismo modo que hay grados diferentes de antirrealismo, hay distintos grados de realismo. El realista que mantiene *que podemos* concebir que ciertas proposiciones son verdaderas de un modo irreconocible, debe contestar a la misma pregunta sobre qué quiere decir «irreconocible». Pero no nos detendremos en esta complicación añadida.

Podemos volver ahora a la cuestión de si el coherentismo es una forma de antirrealismo. ¿Qué tipos de afinidad existen entre las formas de antirrealismo y de coherentismo? Una posición clásica en filosofía es la combinación de idealismo y coherentismo, no sólo como teoría de la justificación, sino también como teoría de la verdad. Esta combinación se encuentra en Bradley y Blanshard. Y podemos ver por qué es probable que un idealista sea coherentista. El idealista sostiene que no hay más que experiencia y creencia; incluso podría, como sugerimos en el último capítulo, decir simplemente que no hay nada distinto a la creencia, sensorial o no sensorial. Si eso es así, no hay nada distinto al conjunto de creencias que pudiera utilizarse para dis-

criminar las creencias verdaderas de las falsas. La verdad debe ser una característica interna, una relación entre elementos del mismo tipo, más que una relación externa entre experiencia o creencia y algo diferente. Y el coherentista concibe la verdad como sólo una característica interna. De modo que es posible que un idealista debiera ser coherentista.

Pero todavía no parece existir razón alguna por la que el coherentista haya de ser idealista; el vínculo entre ambos es unidireccional. Podemos concebir la verdad con una relación interna al modo coherentista, sin añadir el punto de vista idealista de que las proposiciones relacionadas de este modo tratan de la experiencia.

Sin embargo, podríamos sentir la tentación de pensar que el coherentista, si no es un idealista, estará obligado a adoptar el antirrealismo puro. Y ello porque en esta posición se encuentran combinadas las teorías de la coherencia de la verdad y de la justificación. Si adoptáramos sólo la teoría de la coherencia de la justificación, combinándola con una teoría de la correspondencia de la verdad, sería fácil ver que la verdad quedaría como algo trascendente a toda evidencia. Puesto que la justificación, por fuerte que sea, sería sólo una relación interna. Y siempre sería posible la presencia de una relación interna y la ausencia de la verdad, la relación externa entre el conjunto coherente y el mundo. Este punto de vista, que constituye una mezcla, es una forma de realismo [un ejemplo del cual puede encontrarse en Ewing (1934), p. 250]. Pero argumentamos en 8.2 que se trata de una combinación poco atractiva, porque es difícil ver cómo, en tal mezcla, una creencia justificada tiene más posibilidades de ser verdadera.

La combinación más consistente de coherentismo en teoría de la verdad y teoría de la justificación acepta la imposibilidad de una verdad que sea trascendente a toda evidencia y, por tanto, se compromete con el antirrealismo. Dado que, a medida que la justificación se incrementa, se acerca a la verdad, no parece haber ninguna separación entre la «mejor justificación posible» y la verdad. Y esto es todo lo que necesitaba el antirrealista para derrotar al argumento del error.

Si esto es correcto, el coherentismo se revela como una forma de antirrealismo y como la base para una epistemología inmune a nuestros argumentos escépticos. Pero si, por otra parte, un realista puede adoptar el coherentismo de forma consistente, perdemos ese resultado tan reconfortante. Y algunas formas de coherentismo parecen compatibles con el realismo. La cuestión gira de manera crucial en torno a la actitud que tomemos respecto a la noción idealizada de un conjunto de proposiciones completamente coherente y sobre el sentido exacto que

demos a la doctrina de los grados de verdad. Los coherentistas no son unánimes respecto a esta doctrina, y esas diferencias constituyen la base de la que surge la distinción entre los que son realistas y los que no lo son.

Los realistas pretenderán que no hay nada como el conjunto de todas las proposiciones verdaderas. Para cada conjunto, existe en un momento dado, o incluso de forma intemporal, un conjunto que es mayor y más coherente. Y, si la verdad de las proposiciones se define en términos de la coherencia de los conjuntos, no puede existir una proposición situada en el mismo límite de la verdad. Ni siquiera podemos decir que las proposiciones son verdaderas de un modo aproximado, o que se acercan más y más a la verdad en tanto que miembros de conjuntos cada vez mayores. Esto sería suponer que existe un estado imaginable que no puede ser mejorado, al que denominamos «ser verdad» y al que nos aproximamos gradualmente. Pero no existe tal estado, se aleja de nosotros a medida que nos aproximamos. Ninguna proposición, pues, es perfectamente verdadera, dado que la idea de verdad perfecta es un ideal inalcanzable.

Desde este punto de vista, parece que queda el espacio suficiente para que el coherentista sea un realista. La verdad, con el *status* de ideal, yace más allá de cualquier posible conjunto de creencias. Cualquier conjunto de creencias semejante, por enorme que sea, podría ser falso. No hay nada como un punto del que pudiera decirse que, una vez alcanzado, ya estamos situados en un lugar en el que no existe la posibilidad de falsedad. Y la razón no está en la brevedad de la vida humana o la falibilidad del sistema cognitivo de los humanos. Se trata, más bien, de la relación entre la finitud de cualquier posible conjunto de creencias y la infinitud del ideal de un conjunto completamente coherente. De acuerdo con este punto de vista sobre la teoría de la coherencia de la verdad, la verdad es trascendente a toda evidencia en un sentido muy fuerte.

Esta conclusión es discutible e importante. Argumentaremos en 11.1 y 11.6 que el antirrealismo se compromete con una teoría de la percepción inadecuada. Si todos los coherentistas fueran antirrealistas (o, por ejemplo, idealistas), esta conclusión destrozaría el coherentismo. De modo que es importante que exista una forma de coherentismo compatible con el realismo. Afortunadamente, parece que es así.

Sea esto correcto o no, debemos ser precavidos a la hora de aceptar como algo obvio que el antirrealismo constituye una posición invulnerable al escepticismo, particularmente al argumento del error. Es verdad que los dos primeros argumentos escépticos del capítulo 1

apelan a la hipótesis inverificable de que somos cerebros en una cubeta a los que se provocan las experiencias que tenemos. Pero el argumento del error, aunque se presenta como dependiendo de la creencia de que la verdad trasciende a toda evidencia, puede funcionar de hecho con relación a la observación mucho más simple de que los seres humanos *nunca* estamos en una posición tan buena que no haya espacio para el error, para la posible falsedad de nuestra creencia actual, dadas nuestras evidencias. Incluso si la mejor evidencia posible para *p* y la verdad de *p* no están separadas, siempre existirá para nosotros esa separación entre nuestra evidencia y la verdad de *p*. ¿Cómo podemos, entonces, discriminar entre la situación en la que *p* es verdad y aquella en la que *p* es falsa? Independientemente de lo fuerte que pueda ser nuestra antipatía a la existencia de diferencias inverificables, esta situación escéptica todavía puede hacer daño.

La conclusión de este capítulo es la de que el coherentismo no es vulnerable a ningún argumento basado en el regreso de justificaciones, pero, con ello, todavía no tenemos una respuesta al escepticismo.

LECTURAS COMPLEMENTARIAS

La distinción entre internalismo y externalismo sólo se ha convertido en foco de discusión recientemente. Goldman (1980) y Bonjour (1980) discuten los méritos relativos de ambas posturas, aunque no parecen estar de acuerdo en qué sea el internalismo.

Pollock (1970) apoya una forma de internalismo, como hace Kvanvig (1984).

Formulaciones clásicas del idealismo y del fenomenalismo son las de los *Principios del Conocimiento Humano* de Berkeley (*Principles of Human Knowledge*, edición de 1954 o cualquier otra fiable) y Mili (1867, pp. 244-257).

Como contraste, Dummett (1978, cap. 10) constituye una introducción al antirrealismo puro.

Los intentos de Kant de combinar idealismo y realismo se discuten en Stroud (1984, cap. 4) con referencia especial al éxito de Kant en contra del escéptico.

El externalismo en filosofía de la mente se discute por diversos autores en el colectivo de Woodfield (1982), sin excesivas referencias a sus efectos en epistemología.

PARTE III

FORMAS DE CONOCIMIENTO

10. TEORÍAS DE LA PERCEPCIÓN

10.1. ¿HAY ESPACIO PARA LA FILOSOFÍA DE LA PERCEPCIÓN?

Hay dos cuestiones generales que debemos afrontar antes de entrar en el intrincado territorio de la teoría de la percepción. La primera es la de qué pueden esperar añadir los filósofos a los estudios sobre la percepción que llevan a cabo los psicólogos y los neurofisiológos. La segunda es la de cómo un epistemólogo puede conceder el respeto debido a los resultados de esos estudios sin quedar involucrado en alguna forma de circularidad viciosa.

Para responder a la primera cuestión, quizá deberíamos decir que los filósofos están interesados en cuestiones muy generales sobre la percepción. Éstas son cuestiones sobre las que los psicólogos tienen, y están en la obligación de tener, sus propios puntos de vista. Estos puntos de vista generales deben estar de acuerdo con sus descubrimientos más específicos, por lo que no debe esperarse que los filósofos eludan el conocimiento de algunos de esos descubrimientos. Pero las cuestiones más generales pueden ser discutidas de propio derecho, como espero que muestren los capítulos siguientes.

La segunda cuestión la plantea nuestra propia respuesta a la primera. Si un epistemólogo debe aceptar como relevante el tipo de resultados que proporcionan los psicólogos y los neurofisiológos, ¿cómo puede evitar la acusación de que su estrategia es circular? Esta circularidad radica en el intento de construir un análisis general de la creencia justificada que dé por supuesto que las creencias de los psicólogos profesionales, por ejemplo, están justificadas. Para responder a *esta* cuestión necesitamos un análisis de lo que sea la filosofía y de su relación con las ciencias, dado que la afirmación de que los filósofos estudian cuestiones más generales aumenta, más que disminuye, las dificultades respecto a la circularidad. Volveremos sobre este tema en el capítulo 15.

10.2. TEORÍAS DE LA PERCEPCIÓN

Existen tres grupos principales de teorías de la percepción: realismo directo, realismo indirecto y fenomenalismo. En este capítulo haremos un poco de topografía filosófica y seleccionaremos los miembros principales de cada uno de esos grupos, así como el tipo de consideraciones que se usan a su favor y en su contra. Para evitar una confusión excesiva, no atribuiré las posiciones respectivas a sus protagonistas más importantes hasta la sección de referencias al final del capítulo, donde se encuentra también un mapa de las distintas posiciones que puede ser útil.

En primer lugar, para una comprensión inicial de las diferencias entre las tres familias, necesitamos una perspectiva previa de lo que, en este contexto, se quiere decir por «realismo», y del contraste entre el realismo directo y el indirecto.

El realismo en teoría de la percepción (al que denominaré realismo perceptivo) puede caracterizarse inicialmente, de una manera vaga, como el punto de vista de que los objetos que percibimos tienen la posibilidad de existir, y normalmente existen, conservando algunas de las propiedades que percibimos que tienen, incluso cuando no son percibidos. Lo que quiere decir que la existencia de los objetos que percibimos y, al menos, parte de su naturaleza es independiente de la existencia de cualquier perceptor. Este tipo de realismo parece distinto del que discutimos en 1.4 y 9.5, al que denominaremos realismo metafísico. En último término, tendremos que decidir cómo están relacionadas ambas formas de realismo (11.6), pero, por entonces, nuestro análisis del realismo perceptivo ya habrá sido mejorado.

El contraste entre realismo directo e indirecto es notoriamente resbaladizo y, como veremos, de difícil definición precisa. Tristemente, es habitual que los filósofos ataquen el uso ambiguo de este contraste por parte de sus oponentes, cuando ellos mismos no son capaces de hacerlo mucho mejor. La definición inicial que dimos de percepción directa puede resultar vaga, pero no tan vaga como para ser completamente inútil. Podemos considerar que un perceptor P percibe directamente un objeto O, si P percibe O sin percibir ningún intermediario I. P percibirá un intermediario I si, tal y como son las cosas, P percibe O tan sólo en virtud de que percibe I. La relación de P con el intermediario I, si es que existe, no es necesario que sea exactamente la misma que la relación de P con O, y en casos normales no lo será. Sólo se requiere que el modo en que P aprehende I sea análogo al modo en que aprehende O. Casos de percepción indirecta, por ejemplo, serían

la propia percepción en un espejo y la percepción de los actores en la pantalla de la televisión. En ambos casos hay un objeto más inmediato de percepción, al percibir el cual decimos que nos percibimos a nosotros mismos o a los actores; quizá podría decirse que el objeto intermedio, que, en estos casos, aprehendemos de un modo directo, es el reflejo, en un caso, y la imagen de la pantalla, en el otro.

La percepción indirecta se acaba de definir en términos de tipos análogos de aprehensión. Habrá tipos de aprehensión que no sean perceptivos, un posible ejemplo sería la aprehensión del dolor o la aprehensión de la posición de los propios miembros. De modo que, cuando hablamos de percepción como aprehensión, debemos recordar la existencia de aprehensión no perceptiva.

Ahora bien, la disputa entre el realista directo y el indirecto está relacionada con la cuestión de si aprehendemos alguna vez directamente la existencia y la naturaleza de los objetos físicos. Ambos, como realistas, están de acuerdo en que los objetos físicos que vemos y tocamos pueden existir y retener algunas de sus propiedades cuando no son percibidos. Pero el realista indirecto asevera que nunca aprehendemos directamente los objetos físicos; sólo los captamos de un modo indirecto, en virtud de la captación directa de un objeto intermedio (descrito de formas diversas como idea, dato sensorial, percepto o apariencia). El realista directo niega esta pretensión.

El fenomenalismo es una forma de antirrealismo; el fenomenalista niega la existencia de un mundo físico que yazca tras el mundo de experiencia y que pudiera separarse de este último. Para él, el único objeto posible de aprehensión es la experiencia y los complejos de experiencias. No hay realidad alguna aparte de la experiencia. Si tal cosa es así, los únicos objetos de aprehensión son los objetos directos. No hay nada más que pudiéramos percibir de un modo indirecto. Por tanto, el fenomenalista puede ser considerado como alguien que está de acuerdo con el realista directo con relación al carácter inmediato de la percepción y la ausencia de intermediarios en ella, pero de acuerdo también con el realista indirecto en que los objetos directos de percepción no son los objetos físicos. Ésta es una descripción preliminar: será revisada y, me temo, complicada en 10.6.

Necesitamos también decir algo más sobre la noción de «directo» que hemos utilizado. La noción es vaga, básicamente porque no le hemos dado un sentido preciso a la noción correlativa de intermediario; sólo hemos hablado en líneas generales sobre la percepción de una cosa «en virtud de» percibir otra, y sobre la necesidad de que los dos tipos de percepción sean «análogos». Estos términos pueden y deben

refinarse más para poder estar seguros de que hemos captado claramente la distinción entre realismo directo y realismo indirecto. Mientras tanto, el hecho de que la distinción sea vaga no puede utilizarse como argumento en favor de ninguna de las partes.

Nuestra definición de «directo», aunque insatisfactoria, no carece de fuerza. Hay dos nociones cruciales que tienden o bien a inferirse inválidamente de ella, o bien a confundirse con ella. La primera es la noción de infalibilidad. Podemos preguntarnos cómo es posible equivocarse sobre la existencia o naturaleza de algo que aprehendemos directamente. Si el objeto se me presenta de un modo directo, en la forma más íntima posible, no parece que quede lugar alguno para que me equivoque sobre él. Pero esto es un simple error. En el sentido definido de «directo», podemos estar equivocados sobre un objeto incluso cuando lo percibimos directamente.

También podemos ver que se trata de un error si consideramos el argumento de 4.2 al efecto de que nuestro conocimiento de los propios estados sensoriales es falible. No importa ahora si el argumento tuvo o no éxito. Lo que importa es que, aunque hubiera tenido éxito, no planteaba duda alguna sobre el carácter directo de nuestra aprehensión de nuestros propios estados sensoriales. No había tentación alguna de concluir, a partir de nuestra falibilidad en este punto, que nuestra aprehensión es sólo indirecta, y que existe algún otro objeto más directo que nuestros propios estados sensoriales. De modo que no es necesario que la aprehensión directa sea infalible.

Es importante subrayar este punto por las resistencias que puede provocar. Si partimos de la idea de percepción directa, tendemos a pensar en un objeto directamente percibido como algo que se nos muestra abierto y desnudo para que lo inspeccionemos, de tal modo que será imposible que cometamos errores perceptivos sobre su naturaleza e, incluso, de tal modo que no podrán haber hechos sobre él que dejemos de advertir; así que, a este respecto, seríamos no sólo infalibles, sino omniscientes. Si éste es el modelo de percepción directa, es fácil ver que no puede haber una percepción directa del mundo físico y el realista indirecto vence en la disputa por falta de adversario. Pero, en el modelo que estamos considerando, comenzando por la noción de percepción indirecta tal y como ha sido definida, el perceptor directo no necesita ser ni infalible ni omnisciente, por lo que sí se genera un debate interesante entre dos posibles teorías.

El segundo error es suponer que es obvio que un objeto que aprehendemos directamente debe existir y tener las cualidades que le atribuimos en el momento en que lo aprehendemos. Ésta es una idea

sobre el tiempo; la aprehensión y su objeto deben existir al mismo tiempo. En la formulación precedente de esta idea subsiste cierta sugerencia en favor de la infalibilidad. Pero, olvidándola como algo irrelevante, todavía podemos preguntarnos si el objeto de aprehensión directa no debe existir al menos en el momento de la aprehensión. Ciertamente, el objeto directo se nos presenta, y está presente a nosotros, de un modo directo. Pero ¿cómo puede sernos presente si no es coexistente en el mismo tiempo? En el mejor de los casos, lo no existente debe concebirse como ausente (temporal y espacialmente) y, por ello, no presente.

Hay aquí una ambigüedad desafortunada en la noción de «presente», y mi sugerencia es la de que el argumento anterior se basa en ella. En un sentido de «presente», lo presente se contrasta con lo ausente (temporal y espacialmente). En otro sentido, lo presente es aquello que se nos presenta, aquello de lo que tenemos aprehensión directa. En la medida en que no nos aprovechamos de esta ambigüedad, no veo ningún argumento que pueda llevarnos de un sentido a otro. En el sentido definido, sugiero que un objeto, como una estrella lejana, puede haber dejado de existir en el momento en que lo percibimos directamente.

Ésta es una sugerencia que puede parecer contraintuitiva. Si el lector cree tal cosa, espero que se verá reafirmado cuando advierta que habrá un posible sentido en que la memoria puede ser aprehensión directa del pasado. Menciono este extremo para que no se me acuse más tarde de hacer trampa. No me comprometo todavía con que la memoria sea de hecho una aprehensión directa del pasado; sólo que no podemos argumentar que la memoria ha de ser una forma de aprehensión indirecta sólo porque su objeto haya pasado.

10.3. REALISMO DIRECTO

El realismo directo mantiene que en la percepción sensorial captamos de un modo directo la existencia y naturaleza del mundo físico circundante. Todos los realistas directos están de acuerdo en este aspecto de las cosas: su carácter de objetos directos. Se diferencian, sin embargo, en el grado del realismo que están dispuestos a defender. El realista, en el sentido presente, sostiene que los objetos físicos pueden existir y retener al menos algunas de las propiedades que percibimos que tienen, incluso cuando no los percibamos. La expresión crucial es «al menos algunas», y la cuestión es exactamente cuáles.

Deberemos distinguir dos tipos de realismo directo, el ingenuo y el científico; sin embargo, por la propia naturaleza del problema, deberá haber muchas posibles posiciones intermedias.

El realista directo ingenuo sostiene que los objetos no percibidos pueden retener propiedades de todos los tipos que percibimos que tienen. Con esto quiere decir que un objeto no percibido todavía puede tener no sólo forma y tamaño, sino también el calor y el frío, el color, el sabor y el olor, la aspereza y la suavidad, el ser silencioso o el hacer ruido. La ingenuidad de esta posición radica en la palabra «todas». La posición se convierte en algo menos ingenuo en la medida en que «todas» se convierte en «casi todas», luego en «la mayoría», y así sucesivamente; pero nos es más fácil considerarla en su forma más extrema y descamada.

El realismo científico directo se opone a la forma ingenua de realismo directo. Esta versión científica cree que la ciencia ha demostrado que los objetos físicos no retienen, cuando no son percibidos, todas las propiedades que parecen tener cuando los percibimos. La existencia de algunas de esas propiedades depende de un perceptor. Así, el color, el sabor, el sonido, el olor, el calor y la aspereza no son propiedades independientes del objeto que puedan conservarse cuando nadie lo percibe. El objeto sólo las tiene con relación a un perceptor. El realista directo científico acepta el carácter directo de nuestra percepción del mundo, pero restringe su realismo a un grupo especial de propiedades.

Esta distinción es un pariente cercano de la distinción de Locke entre cualidades primarias y secundarias (véase J. Locke, edición de 1961, libro 2, cap. 8). Locke mantuvo que las cualidades primarias de forma, tamaño, textura molecular y movimiento tienen un *status* diferente del de las cualidades secundarias como el color, el calor, el olor, el sabor, etc. (podríamos denominar a estas propiedades «sensoriales»). De acuerdo con Locke, un objeto que percibimos como coloreado no retiene, cuando no lo percibimos, ninguna propiedad de ese tipo tal y como se nos da en percepción. Por supuesto, hay un sentido en el que se puede decir que todavía tiene color, dado que las propiedades primarias continúan estando presentes de tal manera que aparecerá como coloreado a un perceptor, en las circunstancias apropiadas. Un objeto retiene la «base de cualidades primarias» para las cualidades secundarias que parece tener. Pero el color–tal–como–lo–vemos, el calor–tal–como–lo–sentimos o el sabor–tal–como–lo–saboreamos no son propiedades de las que pueda decirse que el objeto las conserva cuando nadie lo percibe, y, por esa razón, no podemos suponer que son

propiedades independientes del objeto cuando sí lo percibimos. El color–tal–como–lo–vemos es más una propiedad de nuestra manera de aprehender el mundo que una propiedad del mundo mismo.

Locke proporcionó un conjunto de argumentos para su posición. De formas diversas sugirió que la evidencia experimental es abrumadora y que la alternativa ingenua es inconcebible. Un argumento más débil que la apelación a la inconcebibilidad de una alternativa es uno análogo a la navaja de Ockham. Esta navaja es el principio general de que no debemos multiplicar las entidades más allá de lo que es necesario; por ejemplo, no debemos admitir la existencia de números o de otros supuestos tipos de entidades a no ser que no tengamos literalmente ninguna alternativa. El análogo a este argumento en el caso que nos ocupa es que no debemos admitir la existencia de un tipo de propiedad si lo podemos evitar. Es posible que la parsimonia sea la actitud correcta en metafísica y, además, podemos poner en juego la apelación a la ciencia. Podemos decir que el tipo de explicación que nos ofrece la física actual no necesita suponer que las cualidades secundarias sean propiedades independientes de los objetos físicos. Esto puede mantenerse de dos maneras distintas. En primer lugar, necesitamos apelar a las cualidades primarias de los objetos microscópicos para explicar las cualidades primarias de los macroscópicos, pero no necesitamos adscribir cualidades secundarias a los objetos microscópicos para explicar las cualidades secundarias de los macroscópicos. La forma y el tamaño de un objeto ordinario se explican por las formas y tamaños de sus partes componentes, pero el color y el calor de un objeto ordinario no se explican por medio del color y el calor de sus partes. No necesitamos introducir en nuestra explicación el supuesto de que sus partes tengan color o calor en modo alguno. En segundo lugar, nuestra explicación del suceso perceptivo en el que un objeto parece cuadrado es una explicación causal que requiere que atribuyamos a la causa algún tipo de forma (al menos, en casos normales). Pero nuestra explicación de que un objeto parezca azul, en la medida en que sea posible por el momento, no necesita de una propiedad similar al color en el objeto; la explicación depende sólo de ciertas relaciones entre las propiedades primarias del objeto, del ojo, del cerebro y de las condiciones locales. Por tanto, el principio de parsimonia exige que abandonemos la idea de que hay en el objeto algo semejante a los colores–tal–como–los–vemos. El mundo, en la medida en que existe con independencia de los perceptores, es tan sólo un mundo de cualidades primarias.

Esta apelación a la autoridad de la física es realmente impresio-

nante. No hace imposible que el realista directo mantenga su posición, dado que la navaja de Ockham es tan sólo un principio metodológico y no una verdad necesaria. Pero sí plantea un problema. Creo que la forma ingenua de realismo directo es, a pesar de todo, más atractiva que la forma científica. ¿Cuáles pueden ser las razones de esa preferencia?

Hay dos observaciones que pueden hacerse en contra de la forma científica. La primera es una desazón general sobre el funcionamiento de la distinción entre cualidades primarias y secundarias, que planteó en primer lugar Berkeley (*Principles*..., edición de 1954, §§ 8, 14-15; véase también el primero de sus *Three Dialogues*). La cuestión es la de si nuestra experiencia del mundo, una experiencia en la que las propiedades primarias y secundarias aparecen entremezcladas y con un *status* semejante, es de tal tipo que podemos concebir la clase de separación que requiere la perspectiva científica, y podemos suponer que comprendemos qué sería un mundo espacial con cualidades primarias pero sin las secundarias. La cuestión, por tanto, es la de la inteligibilidad de la concepción científica del mundo. No podemos imaginar qué sería percibir ese mundo tal y como es en sí mismo. Sería imposible ver los objetos de ese mundo y distinguirlos entre sí sin percibirlos, al mismo tiempo, como coloreados. ¿Qué nos convence, pues, de que podemos concebir qué sería ese mundo sin poder concebir cómo parecería? Se diría que hay, en este punto, el peligro de atribuirnos capacidades a las que no podemos dar demasiado sentido.

La primera observación es una desazón general sobre la distinción entre cualidades primarias y secundarias. La segunda observación es la de que, incluso aunque esta distinción fuera impecable, no es obvio que resulte accesible al realista directo. Puede haber una incompatibilidad de perspectivas a este respecto. Si la hay, la distinción no sería fácil de establecer. Una sugerencia inicial sería que, para el realista directo, nuestra aprehensión del calor o del color difícilmente puede ser de orden distinto a nuestra aprehensión de la forma o el tamaño. Ambos tipos de propiedades se nos presentan de un modo igualmente directo; el color de un objeto es tanto una parte de él como lo puedan ser su forma o su tamaño. Una segunda versión podría ser la de que no está claro qué análisis de la percepción de colores puede proporcionar el realista directo científico. Existe el peligro de que su análisis acabe siendo una forma de realismo indirecto, dado que, seguramente, el color se percibe tan directamente como cualquier otra cosa, por lo que, si los objetos físicos no fueran coloreados en sí mismos, deberíamos descubrir algún otro tipo de objeto intermediario que *sí* fuera colorea-

do. Y admitir un objeto intermediario como portador de propiedades sensibles abriría la puerta al abandono completo del realismo directo.

Estas observaciones son un tanto oscuras, pero expresan una dificultad para el realista directo científico y tratan de mostrar que la forma ingenua de realismo es más consistente. La forma ingenua también tiene sus problemas, pero veremos que la mayoría de ellos son, asimismo, problemas para la forma científica.

El primer problema lo constituye la inclinación a rechazar la posibilidad de concebir los colores–tal–como–los–vemos como algo que existe cuando nadie los percibe. Si esta observación fuera correcta, sería una crítica tanto a la forma ingenua como a la científica, dado que la variante científica considera que tal cosa es concebible, aunque la ciencia nos proporciona razones para creer que, por más que concebible, no sucede de hecho. Pero ¿está justificado el rechazo? (Y, si lo está, ¿qué impide que sea igualmente efectivo en el caso de la forma-como-la-vemos y el tamaño–como–lo–vemos, tal y como argumentó Berkeley?) No podemos apoyarnos en nuestras intuiciones básicas, dado que éstas, en la medida en que las haya, me parecen que están todas del lado del realista ingenuo. ¿Qué argumento puede haber tras ese rechazo?

El primero de los argumentos es que el color percibido se altera de acuerdo con el estado no sólo del objeto, sino del medio circundante (básicamente, la luz) y del perceptor. Por tanto, no puede haber tal cosa como el color real del objeto; cualquier elección entre los diversos colores percibidos será arbitraria y dependerá, probablemente, de propósitos de los seres humanos. Y, si el color real se define en términos de los propósitos humanos, ¿no abandonamos la pretensión de que los objetos conservan el color cuando nadie los percibe? La respuesta es no. No es preciso que el color *real* sea el que se conserva cuando nadie lo percibe; en vez de eso, haríamos mejor en definir el color real como el que se percibe en condiciones normales. Esto deja abierta la posibilidad de que los objetos conservan un color no percibido, que puede o no ser su color real. Al mantener que conservan un color no percibido, el realista ingenuo no se compromete a decir qué color es. De cualquier modo, el argumento que parte de los cambios en el color percibido causados por cambios en el observador no es válido. La sugerencia de que algún color se mantiene no percibido es compatible con el punto de vista de que percibimos objetos físicos directamente como coloreados y con el punto de vista de que el color de un objeto puede alterarse de acuerdo con el medio circundante (por ejemplo, una luz de neón o la del anochecer).

La segunda objeción es que los objetos pierden su color en la oscuridad. Por tanto, puede decirse que el color sólo existe en condiciones adecuadas a la percepción. Es esto lo que hace difícil afirmar que los colores existen sin ser percibidos. Pero, en primer lugar, el realista ingenuo no necesita aceptar que los objetos pierden su color en la oscuridad; puede decir, en vez de ello, que retienen un color que, a causa de condiciones contingentes desfavorables, no podemos ver. E, incluso aunque admita que el color desaparece en la oscuridad, todavía tiene una escapatoria, dado que el hecho de que el color sólo puede existir en presencia de la luz adecuada no contribuye en nada a mostrar que el color no puede existir sin ser percibido, en la medida en que admitamos que la luz adecuada puede estar presente sin ser percibida.

El segundo bloque de problemas para el realista ingenuo es más preocupante y trata de la posibilidad del error perceptivo y la alucinación. Hay una manera correcta y una incorrecta de expresar esta dificultad. La incorrecta es decir, como se hace a menudo, que el realismo ingenuo se refuta simplemente por la existencia del error perceptivo, dado que no podemos percibir algo directamente si no resulta ser como vemos que es. Éste es el viejo error de suponer que la aprehensión directa entraña la infalibilidad, y ningún tipo de realista directo debería preocuparse de él. Sin embargo, el mero enunciado de que el carácter directo de la percepción es compatible con el error perceptivo parece insuficiente, puesto que tal error necesita una explicación y no es en absoluto obvio cómo puede proporcionarla el realista directo. Una manera adecuada de expresarlo sería decir que no es probable que el realista directo tenga una explicación para el error perceptivo sin caer en el realismo indirecto. Éste era un movimiento contra el realismo directo científico y su explicación de los colores; aparece ahora, generalizado, como una dificultad para cualquier realista directo.

Me tomo muy en serio esta dificultad. Pero antes de considerarla (en 11.3) necesitamos examinar el realismo indirecto para ver cómo y por qué pretende obtener mejores resultados. Sea como sea, es difícil comprender las pretensiones del realismo directo hasta que veamos qué mantienen sus oponentes.

10.4. REALISMO INDIRECTO

El realismo indirecto mantiene que en la percepción aprehendemos de un modo indirecto los objetos físicos que nos rodean en virtud de la aprehensión directa de objetos internos, objetos no–físicos.

Ésta es la forma más contundente de realismo indirecto, y es la que nos ocupará en la presente sección. Hay otras formas de realismo indirecto que se mencionarán (sólo para ser rechazadas) en el capítulo 11. Incluyen la sugerencia de que nunca *aprehendemos* objetos físicos, sino que inferimos su presencia de los objetos internos que sí aprehendemos directamente.

Por el momento, sin embargo, nos concentraremos en argumentos en favor de la teoría tal y como se ha formulado. La tradición contiene una enorme cantidad de argumentos diseñados para forzarnos a admitir la presencia del objeto interno. Ofrezco, a continuación, cuatro de ellos.

El primero se basa en la introspección. Ciertamente, cada uno de nosotros sabe, a partir de su propio caso, que cuando aprehendemos un objeto externo nuestro estado perceptivo tiene su propia naturaleza. Dos personas que perciban el mismo objeto estarán en estados perceptivos distintos. Y las diferencias entre esos estados pueden ser descritas como diferencias en *contenido*. Estos estados distintos tienen contenidos perceptivos diferentes. ¿Y cómo podría el contenido de nuestra percepción ser algo distinto a un objeto de percepción? Se supone que aprehendemos ese contenido y que sólo en virtud de nuestra aprehensión del mismo podemos aprehender (indirectamente) el objeto material que se nos presenta. Así que el contenido de nuestro estado perceptivo es el objeto directo de aprehensión, el objeto intermediario.

El segundo argumento es conocido como el argumento del intervalo temporal. Cuando percibimos un objeto como una estrella, es probable que la estrella haya dejado de existir en el momento de la percepción, por lo que al no ser contemporánea al acto de la percepción, sólo podría ser el objeto indirecto del mismo. Y, si es el objeto indirecto, deberá existir otro objeto que sea el directo; y, si en este caso existe tal tipo de objeto, también deberá existir en los demás casos, dado que siempre hay un intervalo temporal, por pequeño que sea, entre el momento en que un objeto tiene determinada propiedad y el momento en que percibimos que la tiene. Si esto es así, aunque el objeto no haya cesado de existir, el estado del objeto que se nos presenta en la percepción no existe en el momento del acto perceptivo y no puede constituir, por tanto, el contenido de la percepción, que sí debe estar presente en dicho momento.

Claramente, este argumento descansa, en gran medida, en la segunda de las confusiones que se señalaron en 10.2; la de suponer que el objeto directo de percepción debe estar presente en el momento de la percepción. Lo que añade a esa confusión son consideraciones

como las del primer argumento. La aprehensión debe tener un contenido, el contenido debe existir en el momento de la aprehensión, y debe ser el objeto directo interno. Dado que éste *es* el primer argumento, el argumento del intervalo temporal no parece añadir nada nuevo.

El tercer argumento es el conocido argumento de la ilusión. Ha aparecido de muchas maneras; de hecho, hay un sentido en el que el argumento anterior podría considerarse un argumento de la ilusión (incluso dos sentidos, siendo uno de ellos un tanto peyorativo). Pero la forma central del argumento de la ilusión se basa en el hecho de que las experiencias perceptivas genuinas, en un momento dado, son cualitativamente indistinguibles al perceptor de las experiencias ilusorias. Debemos tener cuidado con la palabra «ilusión» en este contexto. El argumento adopta su forma más fuerte cuando nos concentramos no en la ilusión, cuando un objeto nos parece ser diferente a como realmente es, sino más bien en la alucinación, cuando creemos que hay un objeto frente a nosotros que en realidad no existe. El realista directo parece obligado a aceptar que, en casos normales, la percepción es una relación entre perceptor y objeto (externo), pero que existen casos anormales de alucinación que, aunque indistinguibles al perceptor en su naturaleza general, son de un tipo completamente distinto, constituyendo un estado no relacional del perceptor más bien que una relación entre perceptor y objeto. La semejanza fenoménica entre los dos casos (es decir, el hecho de que la persona implicada no pueda discriminarlos) constituye una razón para evitar cualquier análisis que las suponga de naturaleza radicalmente distinta. Por ello, el realista indirecto concluye que el mejor análisis es suponer que, aunque ambos estados tienen un objeto interno, sólo uno tiene un objeto externo indirecto. Dado que la existencia y la naturaleza del objeto indirecto no es intrínseca a la naturaleza del estado perceptivo del perceptor, las semejanzas fenoménicas entre los estados no impiden que podamos decir que uno está agraciado con un objeto externo del que el otro carece.

El argumento no es concluyente. Sería posible suponer que los dos estados son fundamentalmente diferentes, a pesar de sus diferencias fenoménicas, pero el argumento de la ilusión nos recuerda que hay, como mínimo, algo extraño esta estrategia, y que sería mucho más atractivo evitarla. Estoy de acuerdo en que el realismo directo está comprometido con esa extraña distinción, pero creo que, a pesar de ese carácter extraño, todavía puede ser preferible al realismo indirecto (las razones pueden encontrarse en 11.2).

El argumento final para el realismo indirecto es un argumento

basado en los logros de la neurofisiología. Este argumento pone de relieve los procesos causales involucrados en la percepción, cuyos detalles estamos comenzando a vislumbrar. Dada esa complejidad, nos pregunta: ¿cómo podemos pretender que percibimos directamente el objeto externo? Hay numerosos estados o procesos intermedios en el cerebro entre el objeto externo y la percepción; es cierto que el objeto está separado de nosotros y sólo puede ser percibido indirectamente por medio de los efectos que tiene sobre la superficie de nuestras retinas, etc.

El argumento es un error. Es fácil que estemos de acuerdo sobre la importancia de los recientes descubrimientos neurofisiológicos, pero no es tan fácil el acuerdo sobre su relevancia en este contexto. Exactamente, ¿cuál es el sentido en el que los procesos cerebrales ocurren como intermediarios entre nosotros y los objetos externos? Es un sentido causal; tales procesos son causalmente necesarios para que se dé la percepción. Pero lo relevante es que no somos conscientes de su intervención en ningún sentido remotamente semejante al sentido en el que somos conscientes de los objetos externos. De modo que los procesos neurofisiológicos no funcionan como objetos directos e intermediarios de la percepción. El argumento, pues, no es un argumento para el realismo indirecto, en absoluto.

Lo que nos queda como resultado de estos argumentos no es más que la observación, incluida en el primero, de que el contenido de la aprehensión debe considerarse un objeto interno y el carácter extraño, que puede no ser letal, del realismo directo, tal como lo revelaba el argumento de la ilusión. Ninguno de estos extremos es concluyente. El segundo puede esperar hasta que examinemos con más detalle el éxito, o la falta de éxito, con el que el realista directo puede enfrentarse al fenómeno del error perceptivo (11.3). El primero es más una afirmación que un argumento. Su éxito depende de un examen de las dificultades para considerar los contenidos de la percepción como objetos. Tales dificultades se exploran en 11.2.

10.5. LAS FORMAS INGENUA Y CIENTÍFICA DEL REALISMO INDIRECTO

Podemos distinguir entre las formas ingenuas y científicas del realismo indirecto. La forma ingenua mantiene que el objeto indirecto de aprehensión tiene propiedades de los mismos tipos que las propiedades del objeto directo. De modo que el objeto indirecto, el objeto físi-

co, tiene tanto color, olor, sabor, etc., como forma y tamaño. El realismo científico indirecto, que es mucho más común, sostiene que el objeto indirecto sólo tiene las propiedades primarias, y que las propiedades secundarias (sensoriales) sólo pertenecen al objeto directo. No hay, como diría Locke, nada semejante a ellas en el objeto (indirecto).

Ambas opciones plantean problemas. La forma científica entraña el tipo de separación entre las propiedades primarias y las secundarias que nos cuestionábamos anteriormente. Pero la manera en que la separación funciona en el interior del realismo indirecto es diferente del modo en que funciona en el seno del realismo directo, y podría pensarse que es menos cuestionable. Después de todo, una de las objeciones a ella en 10.3 era su incompatibilidad con el espíritu del realismo directo, y el hecho de que era probable que condujera a algún tipo de realismo indirecto en el análisis de los colores y de otras propiedades sensoriales. Se planteó, pero no se contestó, la cuestión de si podemos concebir un mundo de objetos que sólo posea propiedades primarias. El asunto es discutible. De modo que, por todo lo que se ha dicho hasta ahora, la forma científica de realismo indirecto es una opción posible.

La forma ingenua, sin embargo, es claramente insostenible. Entraña la grotesca sugerencia de que, junto con los colores visibles pertenecientes al objeto directo de aprehensión, hay otros colores que no se pueden ver en el mismo sentido y que pertenecen a los objetos indirectos. No podemos aceptar la sugerencia de que hay dos tipos de color, uno visible y otro invisible; o, para no exagerar sobre los compromisos de una teoría, que hay dos maneras de aprehender el color, una directa y otra esencialmente indirecta. Sentimos la fuerte tentación de decir que el color es algo que sólo podemos aprehender directamente, al menos en los casos normales. Y creo que no deberíamos resistirnos a tal tentación.

En las dos últimas secciones he ofrecido un mapa de los cuatro tipos de realismo en teoría de la percepción. He argumentado que la forma ingenua de realismo directo es más sólida que la forma científica, y que la forma científica de realismo indirecto es más sólida que la ingenua. No hemos encontrado argumentos concluyentes en favor de ninguna de las teorías supervivientes, ni, en líneas más generales, en favor del realismo directo, por contraposición al indirecto (o viceversa). Ofrezco ahora la sugerencia preliminar de que el realismo directo está apoyado por la intuición extraordinariamente fuerte de que el mundo físico yace, en cierto sentido, abierto a la inspección directa; a una inspección que no se hace a través de intermediarios. Por otra parte, los pronunciamientos de la física favorecen el tipo de distinción

entre cualidades primarias y secundarias que parece encuadrarse mejor en el seno del realismo indirecto.

10.6. FENOMENALISMO E IDEALISMO

El análisis del fenomenalismo de 10.2 era preliminar; necesita de ciertas complicaciones para comprender propiamente la relación entre fenomenalismo y realismo perceptual.

En 9.5 distinguimos entre idealismo y fenomenalismo como formas de antirrealismo. El contraste inicial entre los dos puede verse más claramente en la obra del obispo Berkeley (*Principles of Human Knowledge*, edición de 1954). Su idealismo era del tipo que podríamos denominar metafísico; pretendía que los objetos físicos sólo son conjuntos de ideas reales. Para él, un objeto no puede existir sin ser percibido; el *esse* de un objeto físico es *percipi* (existir es ser percibido). Es posible que los objetos reaparezcan después de un intervalo en la percepción, pero no pueden continuar existiendo cuando nadie los percibe durante el intervalo. El peligro de que, de acuerdo con este análisis, la mayoría de los objetos serían intermitentes, y los problemas subsiguientes sobre reidentificación tras un intervalo, fueron neutralizados por Berkeley al afirmar que Dios es el perceptor permanente de todas las ideas posibles. Con esta afirmación, se asegura que los objetos físicos tengan una existencia continuada análoga al tipo de existencia que el realista reclamaría para ellos. No va más allá de una analogía, dado que, para el realista, los objetos son continuos e independientes, mientras que, para el idealismo berkeleyano, son continuos y dependientes.

También hay indicios en Berkeley del punto de vista más flexible de que los objetos existen en la medida en que *sea posible* percibirlos, incluso aunque no sean percibidos de hecho (cf. *Principles*, edición de 1954, §§ 3 y 58). Una manera de expresar este punto de vista es decir que, de acuerdo con él, los objetos son conjuntos de ideas posibles o reales; pero esto es un tanto extraño, porque no está claro qué es una idea posible, ni si puede ser reunida en el mismo conjunto con las ideas reales. Sea cual sea el modo escogido para expresar este punto de vista, implica claramente que los objetos físicos pueden existir no percibidos, en la medida en que todavía puedan ser percibidos. Ésta es una conclusión fenomenalista.

El fenomenalismo es más flexible y (por tanto) más verosímil que el idealismo, dado que parece que da un sentido más natural a la idea

de que los objetos físicos no cesan de existir cuando no son percibidos, y, por esa razón, proporciona una teoría de la percepción más adecuada. A diferencia del idealismo, el fenomenalismo puede ofrecer algo a modo de explicación del hecho de que se den sucesos perceptivos. El fenomenalista puede decir que lo que implica el hecho de que yo vea mi coche en el garaje cuando abro la puerta es que el coche esté allí durante todo el tiempo, a la espera de ser visto. (Consideraremos en 11.1 si esta explicación es tan efectiva como parece.) Por esta razón, se considera el fenomenalismo como la forma más importante de antirrealismo perceptual. El idealismo queda fuera del análisis, a pesar de que, en último término, pueda ser una teoría más consistente y de que Berkeley podría haber estado en lo cierto al preferirla.

Hay una distinción entre tipos de fenomenalismo, que es relevante para la manera de concebir la relación entre fenomenalismo y realismo. Se trata de la distinción entre fenomenalismo eliminativo y reductivo. Y la misma distinción puede trazarse entre diversas formas de idealismo, donde puede ser entendida más fácilmente.

La cuestión crucial aquí es la de si el idealista admite o niega la existencia de objetos físicos o materiales. El idealista eliminativo mantiene que no existe tal cosa como un objeto físico, que no hay nada más que experiencia (idea, sensación). El idealista reductive sostiene que sí hay objetos materiales, pero que no son nada más que complejos de experiencias. Berkeley era de este último tipo, aunque complicó la cuestión al distinguir entre lo que querían decir los filósofos (científicos) por «objeto material», que le resultaba incomprensible, y lo que quieren decir las personas ordinarias, que, según él, era verdadero.

La distinción entre fenomenalismo eliminativo y reductive implica que nuestro análisis original del realismo perceptivo necesita ciertas reformas. Dijimos que el realista perceptivo sostiene que los objetos materiales existen (o pueden existir) y retienen la mayor parte de sus propiedades cuando nadie los percibe. Con toda certeza, el fenomenalista eliminativo lo niega. Pero el fenomenalista reductivo lo puede afirmar y lo afirma. Por tanto, necesitamos una base nueva y mejor para la distinción entre realismo perceptivo y sus oponentes. La podemos encontrar cuando consideremos qué análisis dan las teorías rivales sobre la naturaleza del mundo cuando no se experimenta.

Todos están de acuerdo, quizá, en que durante los intervalos entre la experiencia la posibilidad de experiencia se mantiene. Una consecuencia de la existencia permanente de los objetos físicos es la de que, si alguien estuviera en el momento y lugar adecuados, tendría una

experiencia de determinado tipo (o, mejor, el mundo le parecería ser de determinada manera). Pero el realista se distingue por su deseo de afirmar que, tanto como la posibilidad permanente de experiencia, está la base permanente de tal posibilidad, algo distinto de ella y que le da soporte; y este algo es el mundo material. Esta proposición es negada por todos los fenomenalistas y podemos exponer la debilidad del fenomenalismo, precisamente, a partir de esta negativa.

LECTURAS COMPLEMENTARIAS

Ayer (1969, caps. 1-2) proporciona un análisis clásico del argumento de la ilusión, pero extrae de él la conclusión idiosincrásica de que el realismo indirecto no es una teoría sustancial sobre la naturaleza de la percepción, sino una recomendación para un nuevo lenguaje filosófico (de acuerdo con sus puntos de vista sobre la naturaleza de la filosofía).

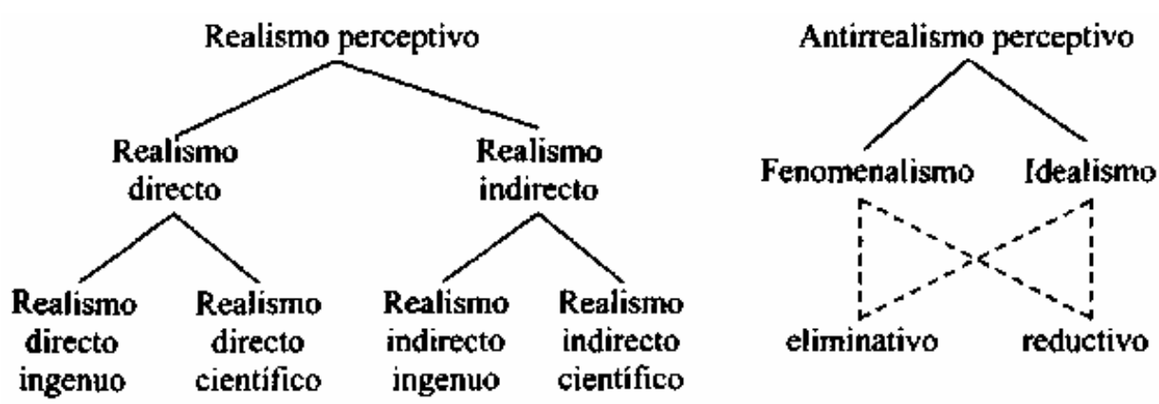
Strawson (1979) apoya el realismo directo ingenuo, pero, de una forma extraña, mantiene que es compatible con alguna forma de realismo científico.

Comman (1975, introducción) proporciona un análisis valioso de la noción crítica de percepción directa; en el capítulo 6 se defiende el realismo ingenuo en contra de la mayoría de las objeciones. Mi discusión debe mucho a la de Cornman.

Jackson (1977, cap. 1) discute la distinción entre objetos de percepción mediatos e inmediatos.

Según mi punto de vista, los libros de Cornman y Jackson son los mejores de entre los accesibles sobre la teoría de la percepción.

He intentado asegurar que mi terminología de este capítulo sea razonablemente representativa. La cartografía de las diversas teorías puede ser la siguiente:



El realismo directo se denomina muchas veces simplemente realismo ingenuo, pero con ello se oscurece una distinción importante. El realismo indirecto se denomina a menudo realismo representacional o teoría representacional de la percepción.

Las diversas teorías se discuten del siguiente modo en este capítulo y en los próximos:

El realismo directo ingenuo se prefiere provisionalmente al realismo científico directo en 10.3.

El realismo indirecto ingenuo se rechaza en 10.5.

El idealismo se rechaza en 10.6.

El fenomenalismo se rechaza en 11.1.

El realismo científico indirecto se rechaza en 11.2.

El realismo directo se examina con más detalle en 11.3-5.

El realismo directo científico se prefiere, sin demasiada seguridad, al realismo directo ingenuo en 11.6.

Los filósofos que pueden considerarse actualmente como defensores representativos de las diversas teorías son:

El realismo directo ingenuo: Strawson (1979), Armstrong [1961; pero cf. Armstrong (1968), caps. 10-11, esp. p. 277], Cornman (1975).

El realismo directo científico: Sellars (1963, cap. 3).

El realismo indirecto ingenuo: Mackie (1976, cap. 2). Se trata de un examen de los puntos de vista de Locke.

El realismo indirecto científico: J. Locke (1961), Jackson (1977).

El fenomenalismo: Ayer (1969, cap. 5), C. I. Lewis (1946, cap. 7).

Durante el período de 1900-1950, la mayoría de los filósofos fueron realistas indirectos. Ahora la balanza se decanta a favor del realismo directo; Jackson (1977) es una excepción a la regla.

La sugerencia que Sellars es lo que he denominado realista directo científico es discutible; Cornman piensa en Sellars como en un realista indirecto [Cornman (1975), pp. 271-272]. El problema radica en si los *senso*, que Sellars concibe como constructos teóricos explicativos, se conciben también como objetos intermedios de percepción.

Aunque Cornman rechaza una teoría que denomina realismo directo ingenuo, su «realismo compatible del sentido común» parece una versión de lo que he denominado realismo ingenuo directo porque es una forma de realismo directo que rechaza la distinción entre cualidades primarias y secundarias.

11. PERCEPCIÓN: LA ELECCIÓN DE UNA TEORÍA

11.1. EL FENOMENALISMO Y LA EXPLICACIÓN DE LA EXPERIENCIA

Supongamos que, en cierta ocasión, le parece al lector que hay una pared enfrente de él, y que está en lo cierto; hay una pared y él la puede ver. ¿Qué explicación puede darse del hecho de que le parece que está enfrente de la pared? ¿Cómo explicar que suceda una determinada experiencia perceptiva?

El realista tiene una respuesta con enorme atractivo, y la cuestión es la de si podemos o debemos resistirnos a su poder de atracción. La explicación parte de la existencia continuada de un objeto material con ciertas propiedades y del suceso nuevo consistente en que un perceptor se ponga en contacto (aunque sólo sea visual) con dicho objeto. La pared estaba ahí todo el tiempo, lo que explica por qué le parece a alguien, cuando llega y abre sus ojos, que ve una pared.

La respuesta paralela del fenomenalista es la de aceptar que hay algo continuo en este caso, lo que explica que, en un momento dado, se dé determinada experiencia perceptiva, pero no se trata de la pared del realista. Hay una continua o permanente posibilidad de experiencia; experiencia que se desencadena cuando se dan las condiciones apropiadas (lo que el realista describiría de nuevo como el hecho de que alguien llegue a la escena con los ojos abiertos). De modo que el fenomenalista explica el que le parezca a alguien que ve una pared apelando a un condicional subjuntivo permanentemente verdadero: si se dieran las condiciones adecuadas le parecería a la persona en cuestión que ve una pared.

El problema es que este condicional subjuntivo parece requerir algún tipo de explicación. ¿Cómo llega a suceder que el condicional subjuntivo sea verdadero? Por supuesto que el realista no negará que sea verdadero, ni siquiera que sea relevante para la explicación. Pero puede proporcionar una explicación de su verdad como lo hacía anteriormente: es verdadero porque existe un objeto físico continuo que actúa *como fundamento* de la posibilidad permanente de sensación. Le

podría parecer a alguien que ve una pared porque (entre otras cosas) hay una pared ahí durante todo el tiempo. ¿A qué puede apelar el fenomenalista para fundamentar el condicional subjuntivo?

La respuesta más común del fenomenalista consiste en apelar a regularidades en experiencias pasadas. El condicional subjuntivo se fundamenta en conjunciones regulares de experiencias de estar en cierto lugar con experiencias de una pared; a alguien le parecería ver una pared en esas circunstancias porque regularmente en el pasado a él (y a otros) le ha sucedido tal cosa. En casos adecuados, podemos inferir un condicional subjuntivo a partir de un enunciado de tal regularidad en la experiencia, y esto es así, precisamente, porque la regularidad convierte al condicional subjuntivo en verdadero.

Pero esta dependencia en las regularidades pasadas como fundamento del condicional subjuntivo parece proporcionar un tipo de explicación incorrecto para la experiencia perceptiva con la que comenzábamos. Podemos verlo mejor con la ayuda de un ejemplo. Supongamos que hay un arco sobre el que descansa la pared. ¿Qué explicaría la capacidad que tiene el arco de servir de soporte a la pared? Una respuesta que apelara al éxito de arcos similares para soportar paredes similares sería incorrecta. Podría acallarnos o dejarnos satisfechos, pero no debería. La petición de una explicación de la capacidad del arco de soportar la pared no se satisface señalando otros casos similares. Sabemos que el arco soporta *de hecho* la pared. Lo que nos interesa es saber por qué. La apelación a las regularidades pasadas no parece decirnos tal por qué, sino sólo que el arco soporta o soportará la pared. Ésta puede ser una información interesante, pero no es la que estamos buscando. Lo que deseamos saber es qué hay en *este* arco que le hace capaz de soportar *esta* pared, y las observaciones sobre otros arcos no parece que sean directamente relevantes, y, de hecho, empeoran el problema. Si hay un problema o misterio sobre *este* arco, hablar sobre otros arcos se limita a hacer el misterio aún más misterioso, no menos.

De un modo parecido, cuando preguntamos por la explicación de una experiencia perceptiva, estamos preguntando por una explicación en términos de *esta* situación; deseamos saber qué hay en esta situación que fundamente su capacidad de producir ciertas experiencias perceptivas. La respuesta del fenomenalista en términos de regularidades previas no es lo que buscamos, aunque sea todo lo que él nos pueda dar. Sólo el realista puede proporcionar una respuesta, en términos de las propiedades inobservadas pero continuas de los objetos que vemos. De modo que el realista puede ofrecer un fundamento contem-

poráneo y relevante para el condicional subjuntivo sobre la experiencia; lo fundamenta en la naturaleza permanente de un tipo de cosa distinto, un objeto físico. Por el contrario, en último término, el fenomenalista no puede proporcionar una explicación de la experiencia perceptiva.

¿Cuál es la relación entre esta crítica al fenomenalismo y la que considerábamos en 6.3? Argumentábamos allí que el significado de un enunciado sobre un objeto material no era equivalente al significado de ningún conjunto de condicionales subjuntivos sobre la experiencia. Y lo que hemos proporcionado aquí es una nueva razón en favor de tal conclusión, dado que uno sólo podría servir de fundamento al otro si no fueran equivalentes en significado. Si fueran equivalentes, no habría ninguna ganancia explicativa en el movimiento de uno al otro. Y un enunciado sobre un objeto material sí puede explicar y fundamentar la verdad de un condicional subjuntivo sobre la experiencia.

¿Qué respuesta daría el fenomenalista a este argumento? La más apropiada sería el cuestionar los supuestos que yacen tras él. Podemos distinguir entre propiedades categóricas y disposicionales de los objetos. Las disposicionales son las capacidades que tiene el objeto para operar en cierto modo, bajo ciertas condiciones. Las categóricas no son disposiciones a actuar de cierto modo (lo que no es todavía decir qué son). Podríamos creer que sabemos que, si un objeto tiene una propiedad disposicional, también debe tener una categórica, dado que las propiedades disposicionales necesitan de un fundamento que no sea disposicional. Por ejemplo, el ser cortante es una propiedad disposicional. Pero un cuchillo cortante debe serlo en virtud de la configuración no disposicional de sus moléculas. Esta configuración es una propiedad categórica que fundamenta la disposición del cuchillo a cortar con facilidad, y es posible que también otras, como la disposición a parecer gris bajo cierta luminosidad.

Podríamos tener la impresión de que las disposiciones no pueden existir sin un fundamento; que, si un objeto tiene la disposición a comportarse de cierta manera o a tener ciertos efectos, es sólo porque tiene ciertas propiedades categóricas, una naturaleza intrínseca que explica y fundamenta su capacidad de comportarse de ese modo. Ésta es la impresión que expresábamos en el argumento de que los condicionales subjuntivos sobre la experiencia necesitan el tipo de fundamento que sólo puede proporcionar un realista, dado que sólo él puede darnos una base no disposicional. Pero esta impresión puede ser cuestionada, a pesar de que para mucha gente sea tan fuerte que la dignifica

hasta el punto de considerarla una necesidad conceptual o un requisito racional. Podemos ver hasta qué punto es cuestionable si advertimos que el tipo de explicación que ofrece normalmente la física actual es disposicional; las propiedades básicas de la materia se conciben actualmente como disposiciones; y, si son disposiciones básicas, es de suponer que no tengan fundamento. Por ejemplo, los físicos conciben la carga eléctrica como una propiedad básica, aunque suponen que hay alguna vinculación entre esta propiedad básica y una disposición a comportarse de cierto modo. Pero ¿cuál es la diferencia entre la propiedad básica y la disposición? No es que una le da soporte o fundamento a la otra; se trata, más bien, de que son indistinguibles. Y, si podemos aceptar propiedades disposicionales básicas en física, seguramente las podremos aceptar en psicología filosófica. Por lo que, en contra del argumento precedente, no es necesario que los condicionales subjuntivos sobre la experiencia tengan un fundamento categórico del tipo que sólo puede proporcionar el realista.

Para contestar a esto, podríamos distinguir tres niveles distintos. Están, en primer lugar, lo que podríamos denominar «disposiciones subjetivas»; se trata de las disposiciones de un objeto a parecer de cierto modo a un perceptor. A continuación las «disposiciones objetivas»: disposiciones a actuar en cierto modo, definibles sin referencia a ningún perceptor. Por último, los estados categóricos del objeto, que fundamentan las disposiciones a actuar de cierta manera, sin ser ellos mismos disposiciones. La respuesta precedente del fenomenalista pretende que las disposiciones subjetivas pueden explicarse sin recurrir a ningún otro tercer nivel de propiedad categórica de las que sólo puede hablar el realista. La respuesta a esto sería constatar que el hecho de que la física considere como básicas ciertas propiedades disposicionales no nos hace más inteligible, en absoluto, la idea de que las disposiciones subjetivas en las que estamos interesados pudieran ser básicas. Cualesquiera que pudieran ser las disposiciones básicas, la disposición a parecer rojo no estaría entre ellas. Así pues, aunque estemos dispuestos a admitir, en algunas ocasiones, la existencia de disposiciones objetivas sin fundamento, no debemos aceptar la posibilidad de un análisis fenomenalista de las disposiciones objetivas. Creo, por tanto, que el movimiento que hace el fenomenalista, aunque mejora en mucho la calidad del argumento, no consigue mostrar cómo sería posible explicar una experiencia perceptiva en términos fenomenalistas.

11.2. REALISMO INDIRECTO: DOBLE APREHENSIÓN Y DOBLE OBJETO

Si se rechaza el fenomenalismo, tenemos que optar entre dos formas de realismo. El resto de este capítulo argumenta en favor del realismo directo. Intento mostrar que no podemos dotar de sentido al papel del objeto «interno» de aprehensión directa, y considero que éste es un argumento suficiente en contra del realismo indirecto. Después, trato de defender el realismo directo de algunas objeciones.

En el capítulo 10 proporcioné los argumentos más importantes en favor del realismo indirecto y de su pretensión de que sólo aprehendemos el mundo físico indirectamente. El realismo indirecto afirma que lo que aprehendemos directamente es algo distinto. La cuestión es qué y cuál sería el papel que podrían desempeñar esos objetos más directos.

Ha habido diversas respuestas a la primera cuestión. Los objetos directos han sido denominados perceptos, datos sensoriales, apariencias, ideas, *sensibilia*, *sema*, etc. La cuestión importante no es, sin embargo, la de qué nombre les demos, sino cómo se supone que son los objetos así denominados. Después de todo, alguien podría decir que ha mostrado que debe haber un objeto directo interno; denominémoslo «dato sensorial» y examinemos con detalle los argumentos tratando de descubrir lo que nos pueden decir sobre la naturaleza de estos datos sensoriales. ¿Qué sentido se desprende de esos argumentos para la palabra «interno», por ejemplo?

El argumento de la ilusión, tal y como lo formulamos en 10.4, nos lleva a suponer que la alucinación es un estado del perceptor que está dirigido a un objeto; pero que, por ser una alucinación, la naturaleza y existencia de ese objeto no dependen lógicamente de ningún hecho relativo al medio circundante al perceptor, ni a la condición física de su cuerpo o sus órganos. La conclusión era que, incluso en los casos normales, podríamos buscar una caracterización de cómo se siente el perceptor, una caracterización que excluyera cualquier detalle sobre el cuerpo del perceptor o sobre el medio circundante. Tal caracterización incluiría un análisis del objeto, un objeto «interno» en el sentido de que su naturaleza dependería por completo del contenido de la percepción, que sería en sí mismo independiente, o podría caracterizarse independientemente, de cualquier implicación sobre el mundo físico. Por lo que, al menos, sabemos que un dato sensorial no puede ser la superficie de un objeto físico, como algunos han sugerido. Pero los argumentos de 10.4 nos dicen muy poco más sobre la naturaleza del objeto directo de aprehensión. Todo lo que sabemos es que se trata de

un objeto portador de las propiedades que conforman el contenido de la percepción, descritas sin ninguna implicación sobre la naturaleza del mundo físico. Creo, sin embargo, que en esta mínima descripción podremos encontrar dificultades para el realismo indirecto. De ellas me ocupo ahora.

LA OBJECIÓN ESCÉPTICA

Una queja tradicional sobre la concepción del realista indirecto es que nos lleva directamente a un escepticismo muy general sobre la posibilidad de conocer el mundo «externo». Si todo lo que percibimos directamente es un objeto «interno», ¿cómo podríamos colocarnos en la situación de adquirir creencias justificadas sobre un mundo físico externo? Por todo lo que sabemos, no existe como mundo físico en absoluto, dado que nunca lo podemos percibir directamente. Berkeley argumentó que no podemos inferir la existencia de un mundo externo. Es lógicamente posible, incluso dada nuestra aprehensión de objetos directos internos, que no haya un mundo externo, por lo que una inferencia de lo interno a lo externo no puede ser deductiva. Por otra parte, si la inferencia fuera inductiva dependería de establecer con éxito correlaciones previas entre los enunciados sobre los objetos directos de aprehensión y los externos e indirectos. Pero, por hipótesis, no podemos establecer tales correlaciones, porque necesitaríamos, para ello, una aprehensión independiente de los objetos externos. El realismo indirecto sostiene, sin embargo, que sólo podemos aprehenderlos de un modo indirecto; nuestra aprehensión de ellos depende de nuestra aprehensión del dato sensorial.

Este argumento ha tenido un efecto enorme, y ha sido uno de los motores más poderosos en el desplazamiento hacia el fenomenalismo. Si alguien está convencido de la existencia de datos sensoriales y abrumado por dudas escépticas sobre cualquier otra cosa, con toda probabilidad, se sentirá atraído por la solución radical de tratar de arreglárselas sin nada más, y de tratar de construir un mundo sobre el único material de los datos sensoriales o conjuntos de tales datos. A pesar de su importancia en el pasado, existe la creciente convicción de que el argumento no está bien concebido. Está equivocado si hace uso de la idea crucial de que no podemos establecer las correlaciones requeridas porque no podemos observar los objetos externos. El realismo indirecto no tiene la consecuencia de que los objetos externos sean inobservables; pretende, simplemente, decirnos algo sobre qué

sea observarlos. Y, si lo esencial es que el realismo indirecto es vulnerable al argumento escéptico del error, la respuesta será que el argumento del error es una amenaza a cualquier teoría de la percepción, incluyendo el realismo directo y el fenomenalismo; dado que, bajo cualquier teoría, siempre existirá la posibilidad de que las cosas no sean como parece que son.

Por ello, la objeción escéptica parece no ser efectiva en contra del realismo indirecto tal y como lo hemos caracterizado. Si el realista indirecto hubiera sostenido que los objetos físicos son inobservables o que inferimos su presencia o su naturaleza a partir de nuestro conocimiento de los objetos internos que percibimos, en ese caso, la objeción escéptica hubiera tenido algún papel relevante. (La importancia de esta observación se verá con más claridad en la siguiente subsección.) Pero contra el realista indirecto que mantenga que hay dos formas análogas de aprehensión, la directa y la indirecta, el argumento escéptico es completamente impotente.

Nos enfrentamos, pues, al problema de si podemos dotar de sentido a la idea de dos formas análogas de aprehensión y de sus objetos correspondientes.

LO DIRECTO Y LO INDIRECTO

La teoría de la percepción de Locke es un ejemplo clásico de realismo indirecto, como hemos caracterizado esa posición. Locke sostuvo que en la percepción hay dos aprehensiones, cada una con su propio objeto. El objeto interno, que denominó idea, es *percibido* y el objeto externo, la cosa material, es *visto*. Así, al percibir una idea en condiciones favorables, *vemos* una cosa material. Hay, por tanto, aprehensión doble: una del objeto interno y otra del externo.

Una objeción natural a la noción de doble aprehensión es que al perceptor sólo le parece que hay *un* tipo de aprehensión y *un* tipo de objeto (podemos discutir más tarde si es o no un objeto interno). ¿Qué se nos puede decir para llevarnos a aceptar la pretensión de que ver un objeto es estar en dos estados de aprehensión distinguibles aunque análogos?

La respuesta es que, de hecho, estamos familiarizados con la idea de que aprehendemos una cosa en virtud de aprehender (análogamente) otra. Aprehendemos la presencia de los objetos que se reflejan en un espejo en virtud de nuestra aprehensión de los reflejos en el espejo, y el sentido en el que vemos los reflejos es análogo al sentido en el

que vemos los objetos en este caso. De un modo semejante, aprehendemos la imagen en una pantalla de televisión en virtud de nuestra aprehensión de las partes que la componen, y aprehendemos los sucesos representados (los vemos sobre la pantalla) en virtud de nuestra aprehensión de las imágenes de la pantalla. Y, de nuevo, aprehendemos (vemos) una vaca en virtud de nuestra aprehensión de la superficie de la vaca que se nos presenta. Así pues, hay ejemplos familiares en los que, dadas las circunstancias adecuadas, aprehender una cosa es aprehender, no del mismo modo exactamente, otra. Por tanto, no es necesario que *parezca* que hay dos formas de aprehensión para que estemos justificados al decir que, en un proceso, se dan dos formas de aprehensión.

Los ejemplos de doble aprehensión difieren entre sí de diversas maneras, lo que no garantiza que alguno de ellos pueda utilizarse como modelo del tipo de doble aprehensión que, de acuerdo con el realista indirecto, se presenta en cada caso de percepción. ¿En cuál de los ejemplos anteriores, si es que sucede en alguno de ellos, se relacionan los dos objetos de aprehensión de un modo análogo a la manera en que el dato sensorial «interno» se supone que está relacionado con el objeto externo?

Los casos de los puntos y la imagen y de la superficie y la vaca son casos en los que los objetos directos de aprehensión constituyen o son partes integrantes del objeto indirecto. No es esto lo que buscamos: los datos sensoriales, de acuerdo con el realista indirecto, no son parte de los objetos externos. Sólo el fenomenalista desearía decir que los datos sensoriales que aprehendemos directamente pasan a ser partes constituyentes de los objetos que vemos, lo que demuestra que, al menos aquí, no nos encontramos con el tipo de caso que el realista indirecto está buscando.

Quizá sean más prometedores los casos de los reflejos en el espejo y los objetos reflejados, y de la imagen en la pantalla y los sucesos representados. En estos ejemplos, la relación entre los objetos directos e indirectos parece ser doble: los objetos indirectos *causan* los directos y se *parecen* a ellos. Esto parece estar mucho más cerca de lo que buscamos, dado que es probable que el realista indirecto afirme, como hizo Locke, que ambas relaciones se dan entre los objetos externos y los datos sensoriales.

Sin embargo, ¿es realmente cierto que nuestros objetos directos internos (si tenemos tal cosa) se parecen a los objetos externos que los causan? Nuestro realista indirecto científico deseará restringir a las cualidades primarias la afirmación de que los datos sensoriales se ase-

mejan a los objetos externos, dado que niega que el color-tal-como-lo-vemos se parezca en modo alguno a la cualidad del color, tal y como se da en el objeto (su base de cualidades primarias). Pero es muy difícil suponer que los datos sensoriales pueden parecerse a los objetos materiales con relación a las cualidades primarias cruciales de forma y de tamaño. No tiene ningún sentido la cuestión de cuánto mide un dato sensorial o una parte de él. Y la impresión de un objeto cuadrado no es, ella misma, algo cuadrado, ni puede pensarse que comparta las propiedades relevantes del objeto cuadrado. Incluso si consideramos datos sensoriales de tipo visual, no podemos suponer que compartan las formas de los objetos que representan. En el mejor de los casos, podemos decir que un dato sensorial puede parecerse a una imagen pictórica de un objeto, pero, entonces, el parecido no se da entre dato sensorial y objeto, sino entre el dato sensorial y una imagen pictórica. El tipo de forma tridimensional que atribuimos a los objetos no es algo que puedan compartir nuestros datos sensoriales. Nuestros datos sensoriales sólo pueden ser semejantes a la manera en que los objetos *parecen*, no a cómo sean. Pero esto no sirve, dado que nuestros datos sensoriales *son* exactamente la manera en que tales *objetos parecen*.

Si lo relevante no es la semejanza, ¿podemos retrotraernos a la relación causal entre objetos directos e indirectos de nuestros ejemplos, para mostrar que son relevantemente similares a la idea del realista indirecto de una doble aprehensión? El problema es que la relación causal es incapaz de sostener la analogía por sí misma. Nuestra aprehensión en estos casos se concibe como algo doble: aprehendemos tanto la imagen como los sucesos representados. Pero esto no es así por el mero hecho de que los sucesos representados estén entre las causas más importantes de las imágenes. Hay muchos sucesos que están entre las causas más importantes de las cosas que aprehendemos, sin que se diga que *aprehendemos* de un modo análogo esos sucesos. En la mayoría de los casos, no los aprehendemos análogamente más de lo que aprehendemos a alguien por ver sólo su sombra. Sabemos que están ahí, pero este conocimiento, aunque lo denominemos aprehensión, no es relevantemente similar a nuestra aprehensión de la sombra. *Vemos* la sombra y no vemos a la persona. De modo que, si abandonamos la semejanza como fundamento de la analogía, destruimos ésta, dado que lo que nos queda es insuficiente para sostenerla por sí mismo.

Las dificultades que hemos encontrado en la idea del realista indirecto de una aprehensión doble no son definitivas, no constituyen una refutación del realismo indirecto. Pero sí nos dan una razón para bus-

car una teoría que no genere tantas dificultades. Una teoría tal sería, por supuesto, el realismo directo; pero, antes de volver a ella, debemos considerar brevemente los méritos o deméritos de otra teoría alternativa.

Esta alternativa, que denominaré realismo inferencial, acepta que hay problemas en la noción de una aprehensión doble, y las evita abandonando una de las dos formas de aprehensión. Pero se diferencia del realismo directo al abandonar no el objeto directo interno o dato sensorial, sino el objeto indirecto externo. La teoría, por tanto, mantiene que no aprehendemos los objetos materiales en ningún sentido análogo a aquel en el que aprehendemos los datos sensoriales. Más bien, nuestro conocimiento de los objetos materiales es *inferencial*. [Un buen ejemplo de realismo inferencial es Russell (1926) y, más recientemente, Ayer (1976).] Se sostiene que los datos sensoriales son los únicos objetos de aprehensión perceptiva, y que, a partir del conocimiento de nuestros propios estados sensoriales, inferimos la presencia y naturaleza de los objetos materiales que los causan.

Una queja sobre este punto de vista es que convierte el mundo material en algo invisible. Comenzábamos tratando de explicar qué es ver, por ejemplo, un objeto material, y acabamos anunciando que tales cosas no pueden ser vistas en absoluto. Como poco, se trata de una consecuencia altamente inverosímil, pero no es letal. Es de suponer que el realista inferencial piense que debemos tragarnos esta medicina porque es nuestra obligación. Pero podemos proseguir con este asunto. Parece que la objeción escéptica, que no era concluyente en contra del realismo indirecto, sí podría serlo en contra del realismo inferencial. ¿Qué razones tenemos para suponer que hay cosas tales como los objetos materiales, si son invisibles, etc.?

La respuesta habitual a este ataque es señalar a generaciones de físicos, cuya práctica ha sido la de inferir lo inobservable a partir de lo observable. ¿No muestra eso que la inferencia de lo observable a lo inobservable no puede descartarse como principio? Pero la analogía es insuficiente. El físico infiere, a partir de la observación de objetos o sucesos estrictamente materiales y normales, un mundo mucho más remoto. La diferencia entre su práctica y la del realista inferencial no es la de que el mundo que infiere el realista inferencial sea remoto, sino la de que los objetos a partir de los cuales infiere no parecen tener las propiedades suficientes como para apoyar la inferencia.

El argumento del capítulo 5 contra la imposibilidad del lenguaje privado proporciona cierto apoyo para esta observación. Si no podemos instituir un lenguaje primario para la descripción de la experien-

cía sensorial, es decir, los datos sensoriales, no podemos esperar sostener la creencia del realista inferencial de que nuestro conocimiento de los objetos materiales se infiere completamente a partir de nuestra aprehensión primaria de datos sensoriales.

De modo que lo que nos resta por hacer es ponderar los méritos del realismo directo. La situación actual es que el realismo directo tiene una enorme plausibilidad intuitiva, pero que parece tener dificultades con el argumento del error. El realismo indirecto tiene menos plausibilidad, y los argumentos en su favor no son concluyentes. Los problemas que hemos encontrado para dar sentido a la idea de una doble aprehensión y un doble objeto parecen, por tanto, suficientes para que escrutemos con mayor detalle el realismo directo.

11.3. EL REALISMO DIRECTO Y LA EXPLICACIÓN DEL ERROR PERCEPTIVO

Al final de 10.3 sugerí que el realismo directo necesita algo más que la afirmación descarnada de que su teoría permite la posibilidad del error perceptivo. Es verdad que la aprehensión directa no entraña infalibilidad. El peligro es, sin embargo, que, al intentar una *explicación* del error perceptivo, el realismo directo acabe por caer en el realismo indirecto.

El atractivo del realismo indirecto en esta área es fácil de ver. Para el realista indirecto, el error perceptivo es un desajuste entre dos objetos; un desajuste, por ejemplo, entre un dato sensorial cuadrado y azul y una caja verde y oblonga. Pero, para el realismo directo, sólo existe un objeto, por lo que le es difícil hablar de desajuste. ¿Entre qué cosas se daría tal desajuste?

El argumento de la ilusión nos pide que admitamos que, en el caso de la alucinación extrema, no está presente, en absoluto, ningún objeto externo. Si es así, ¿cómo analizar lo que le sucede a la persona que sufre la alucinación? Tal persona tiene una experiencia perceptiva, está en cierto estado, perceptivo. Y este estado preceptivo debe poder ser caracterizado de un modo independiente, sin ningún tipo de implicación respecto a la naturaleza del medio físico circundante. Además, tal estado perceptivo, que se da sin el objeto adecuado, puede ser reduplicado en un perceptor que no esté sufriendo ninguna alucinación. En un caso ordinario, también puede darse una descripción independiente del lado de la historia que vive el perceptor, sin decir nada sobre el mundo circundante. Estas descripciones independientes describen un

estado perceptivo. Pero, en el análisis del realista directo, el estado perceptivo no tiene objeto interno y no es un intermediario entre el perceptor y el objeto externo. Parece que, entonces, para que haya un análisis del error perceptivo en términos de desajuste debe establecerse como un desajuste entre el estado perceptivo y el objeto. ¿Cómo ha de ser un estado perceptivo para poder «dejar de ajustarse» al mundo?

Los realistas directos no están obligados a conceder que el argumento de la ilusión haya de conducirlos a este punto. Alguien que adoptara una posición externalista radical en filosofía de la mente (véase 9.3) sostendrá que no hay en la percepción nada como un estado perceptivo cuya naturaleza sea lógicamente independiente de la existencia del objeto del que parece ser una percepción. No hay residuo de tal estado, común a los estados perceptivos y a las alucinaciones, descriptible sin referencia a su objeto. Pero la mayoría de los realistas directos son internalistas en este respecto y admiten la existencia del estado perceptivo como un residuo no–objetual.

¿Qué sentido puede darse a la noción de un desajuste entre el estado perceptivo y el mundo? ¿Cómo son esos estados perceptivos? ¿Qué análisis de ellos podría darse que no los convirtiera en unos nuevos objetos de aprehensión? El realista directo podría apelar, en este punto, a una analogía con la denominada Teoría Adverbial de la Sensación. [Esta teoría se deriva de Chisholm; véase Chisholm (1977), p. 30, y Cornman (1975), pp. 73-77.] Supongamos que estamos preocupados por la expresión «tengo dolor», preguntándonos, quizá, si puede ser verdadera cuando no somos conscientes del dolor. La teoría adverbial intenta evitar concebir el dolor como un *objeto* de aprehensión y reformula la expresión de un modo adverbial: «me siento dolorosamente». Parte del sentido de hacerlo así es sugerir que la expresión «duele» es más reveladora que la expresión «tengo dolor».

La analogía podría usarse para sugerir que debemos analizar el *contenido* de una experiencia visual, o de otro tipo, como una *manera* de ser consciente. En lugar de hablar de experimentar un dato sensorial con una propiedad P, deberíamos hablar de sentirnos P–mente o de sentirnos de un modo P. Incluso si esto equivale a la adscripción de tipos especiales de propiedades a nuestros estados de aprehensión no es incompatible con el realismo directo. Porque el estado de aprehensión que tiene esas propiedades no es un *objeto* de aprehensión, intermedio entre nosotros y el mundo físico. (Ésta es la razón por la que no es probable que Sellars, que mantiene la teoría adverbial, sea un realista indirecto.)

Pero, incluso si el realista directo concibe el estado perceptivo, o el

contenido de la aprehensión, como una manera de aprehender más que como un objeto de aprehensión, ¿de qué le podría servir a la hora de proporcionar un análisis del error perceptivo? Deberá suponer, claramente, que puede haber un desajuste entre el estado perceptivo y el mundo, pero considerar tal estado como una *manera* de aprehender el mundo todavía no proporciona una clara comprensión del posible desajuste entre la manera y el mundo. Todos sabemos que el mundo puede no ser como consideramos que es. Pero tal hecho no proporciona una respuesta a la cuestión. Más bien, *es* la cuestión ¿Cómo podemos entender ese hecho?

Para progresar ahora necesitamos introducir una nueva cuestión sobre la percepción, una cuestión que ya ha aparecido de forma breve e imprevista en 4.2 y 8.4. Podemos considerar la percepción como una forma compleja de sensación; podemos considerarla como, esencialmente, una forma de creencia; o podemos considerarla como una combinación de creencia y sensación. De modo que existen tres tipos de teoría:

1. teoría de la «sensación pura»,
2. teoría «mixta»,
3. teoría de la «pura creencia».

El primer tipo de teoría no necesita sostener que las sensaciones perceptivas son objetos de aprehensión; podría ser una teoría adverbial. Pero es probable ver sólo una diferencia de grado entre sensación «perceptiva» y otras sensaciones más simples, como dolor. Aunque dejemos a un lado la cuestión de si hay sensaciones visuales, distintas, por ejemplo, a las que causa la luz muy brillante e intermitente, todavía nos podemos sentir incómodos con una analogía demasiado fuerte entre la «sensación visual» y una sensación como dolor. Dado que, incluso aunque aceptemos que la percepción involucra el que se dé alguna forma característica de sensación, parece que *ese* tipo de sensación no puede darse sin una creencia, o, al menos, sin que se dé la tendencia a formarse una creencia, sobre los objetos que causan la sensación. La sensación pura no es lo suficientemente cognitiva como para poder constituir un modelo para la percepción. La percepción no es la aparición de una sensación que puede o no provocar una creencia, o, al menos, una tendencia a tener cierta creencia. Percibir es (tener la tendencia a) creer.

Pero esto no nos debe llevar tampoco a una teoría de la percepción como «pura creencia». Hay filósofos [cf. Armstrong (1951), cap. 9]

que mantienen que los elementos «sensoriales» característicos de la percepción no son esenciales al proceso, y analizan la percepción como una mera tendencia a formarse creencias sobre el mundo circundante. De hecho, existe un fenómeno interesante, y descubierto recientemente, que podría considerarse que presta apoyo a esa posición. Se denomina visión ciega. Se da cuando personas que parecen ser ciegas, en el sentido de que no experimentan sensaciones visuales (ésta es, obviamente, una descripción interesada), son capaces de responder con extraordinaria exactitud algunas cuestiones simples sobre la forma y la localización de los objetos circundantes. Las personas en cuestión no saben que sus respuestas son exactas; normalmente, consideran que están tratando de adivinar.

Este fenómeno podría considerarse como una forma de percepción pura no sensorial que mostraría que la sensación visual, aunque habitual, no es algo esencial para la percepción misma. El problema radica en que el tema que nos ocupa en esta investigación sobre la percepción no es esta forma de aprehensión empobrecida, aunque interesante, sino la manera mucho más rica en que nosotros aprehendemos el mundo que nos rodea. La visión ciega no puede utilizarse para mostrar que las diferencias características entre los sentidos, y la diferencia entre ver que algo es verdad y llegar a creer meramente que lo es, no son importantes para *nuestro* modo de escudriñar las cosas. Sea lo que sea la visión ciega, no es ver.

Parece, pues, que debemos optar por una teoría mixta de la percepción, en la que la percepción sea alguna forma de combinación de «sensación» y «creencia». La cuestión es la de qué tipo de combinación podría ser. Lo que tenemos, hasta ahora, es que la percepción ha de ser concebida como algo cognitivo, aunque no podamos dejar a un lado, por ello, las maneras características en las que las cosas le parecen a alguien cuyos oídos están abiertos, etc. Tenemos aquí una elección. O bien consideramos que los dos elementos de la teoría mixta son separables, o bien los consideramos como idénticos en último término. Si los consideramos separables, deberemos suponer que la tendencia a creer que las cosas son del modo en que parecen ser es, de algún modo, extraíble del todo, dejando tras ella la mera apariencia. Si las consideramos como idénticas en último término, deberemos suponer que el hecho de que el mundo se nos parezca en un modo característico *es* exactamente que adquiramos cierta tendencia a creer. Esta tendencia a creer no es algo que pudiera darse en ausencia de ese modo característico de parecer, por lo que no es algo que pudiéramos compartir con el que pueda poseer la visión ciega.

Me parece que esta última alternativa es, con mucho, la más verosímil. En vez de considerar la percepción como una combinación de dos elementos separados, uno cognitivo y otro sensorial, deberemos considerar que la percepción es una *forma* característica de creencia (o de tendencia a creer), una que no comparten los que carezcan de *input* sensorial relevante, y en la que la tendencia a creer no puede separarse de que se dé tal *input*.

Pero, en este punto, en la medida en que optemos o bien por una teoría mixta, o bien por teoría de la percepción como pura creencia, ya podemos ofrecer una respuesta a la cuestión original sobre el error perceptivo y su explicación en términos de desajuste entre el estado perceptivo y el mundo. Este desajuste puede considerarse ahora como un caso del desajuste entre la creencia y el mundo que se da cuando una creencia es falsa. Puede que haya problemas para proporcionar un análisis de la falsa creencia, pero son problemas que, de cualquier modo, estamos obligados a afrontar. De modo que este análisis del error perceptivo, que es accesible al realista directo, no añade problemas nuevos, sino que se limita a unir dos problemas en uno solo.

11.4. UN ELEMENTO CAUSAL

Hemos concedido que una persona que vea, oiga o sienta un objeto, se encuentra en un estado perceptivo que puede describirse sin implicaciones respecto al mundo que la rodea, porque tal estado perceptivo puede darse en alguien, quizá bajo una alucinación, independientemente de cuál sea su medio circundante. Por tanto, no puede decirse que todo aquel que esté en ese estado perceptivo ve (oye, siente) los objetos que lo rodean. ¿Qué más se necesita, además del estado perceptivo, para que la persona implicada *vea?*

No sería suficiente decir que alguien ve lo que lo rodea si y sólo si está en un estado (visual) perceptivo adecuado y el mundo es como él cree, en ese estado, que es. Supongamos que estoy sentado enfrente de una pared blanca, con el cerebro conectado a un computador que me causa diversos estados perceptivos, y que uno de ellos es el de una superficie blanca. ¿Diríamos que hay ciertos momentos en los que veo lo que me rodea y otros momentos en los que no? No lo diríamos, y nuestras reticencias muestran cómo mejorar el análisis de qué estados perceptivos constituyen ver algo. No sólo el mundo debe ser como yo creo, en un estado determinado, que es, sino que cómo sea el mundo debe ser una *causa* de mi creencia de que el mundo es así. Esta idea es

conocida como la teoría causal de la percepción. Ver (oír, sentir) es un estado perceptivo causado de la manera adecuada.

En esta sugerencia causal, hay algo que es obviamente correcto. Me gustaría hacer un par de comentarios sobre ella.

En primer lugar, la sugerencia, tal y como la hemos formulado, no puede ser completamente correcta; la relación causal entre el mundo y el estado perceptivo ha de especificarse en mayor detalle. El ejemplo del científico benevolente revela la necesidad de esos retoques. Supongamos que el computador conectado a mi cerebro está controlado por un científico benevolente que causa los estados que yo tendría si estuviera viendo, es decir, que asegura que mis estados perceptivos no son erróneos (falsos). En este caso, mis estados perceptivos están causados indirectamente por el mundo, pero no diríamos que ésta sería una situación en la que yo *viera* lo que me rodea. La razón parece ser que el vínculo entre el mundo y el estado perceptivo no es fiable, dado que depende de la voluntad de un experimentador que podría ser arbitraria.

Hay, en este caso, un vínculo causal entre el mundo y el estado perceptivo, pero no se trata del vínculo adecuado. ¿Cómo podremos impedir los vínculos causales inapropiados? Una manera de hacerlo sería confiar en los descubrimientos (futuros) de la neurofisiología. Lo que tenemos que hacer es proporcionar una lista de ejemplos que estemos dispuestos a considerar como casos genuinos de ver (oír, sentir) en el medio circundante, y decir que todos los estados perceptivos causados de maneras relevantemente similares a las que están involucradas en los casos específicos pueden ser considerados casos de ver. Esto eliminaría las rutas causales inapropiadas. Lo esencial es que, como filósofos, y/o mortales ordinarios, podemos no tener ni idea de cómo han de ser esas rutas, por lo que dejamos en manos de los neurofisiólogos la investigación pertinente. Y, por el momento, nos servimos por adelantado de sus descubrimientos como si fueran una suerte de pagaré anticipado. [Éste es el punto de vista que adopta Grice (1954).]

Hay, sin embargo, una dificultad en este punto de vista y radica en la expresión «estados causados en maneras relevantemente similares a las que están involucradas en los casos especificados». Supongamos que nuestro pagaré se hace efectivo y que conocemos la historia causal de mi creencia actual con detalle suficiente; ¿conocemos también, por ello, qué otras rutas causales son relevantemente similares y cuáles no lo son? Parece que no es necesario que sepamos tal cosa, por lo que los detalles de la cláusula causal no pueden ser dejados enteramente

en manos de los neurofisiólogos. Si descubren una creencia aparentemente perceptiva causada de una manera completamente nueva, ¿quién podría decidir si la manera era relevantemente semejante a las maneras ya conocidas? Los neurofisiólogos no tienen ninguna autoridad especial para tomar una decisión al respecto. Y nuestra decisión no tendría en cuenta el grado de semejanza entre las historias causales, sino que consideraría directamente nuestra tendencia a aceptar la creencia como perceptiva. Estaríamos decidiendo sobre si queremos incrementar la cantidad de ejemplos cruciales o casos paradigmáticos. Por tanto, la apelación a la neurofisiología parece incapaz de resolver nuestros problemas. Tampoco podemos limitarnos a decir que los estados perceptivos causados de la manera adecuada son casos de *ver*; tal cosa sería eludir la cuestión que nos ocupa. Por tanto, el requisito causal plantea dificultades que, por el momento, no sabemos resolver. [Strawson intentó solucionar estos problemas en Strawson (1974), cap. 4.] Esto nos lleva al segundo de los comentarios que deseo hacer sobre la teoría causal de la percepción.

La necesidad de un elemento causal se mostró al conceder que la diferencia entre alguien que viera su medio circundante y alguien que sufriera una alucinación no debería ser una diferencia «interna». En la medida en que nos limitemos a las personas implicadas, podría no haber ninguna diferencia entre sus estados perceptivos. La diferencia radica por entero en el desajuste en las relaciones causales entre estos estados y el mundo «externo». Este punto de vista nos plantea todas las cuestiones tradicionales como la de qué evidencia «interna» tenemos en un momento dado de estar viendo y no de estar sufriendo una alucinación, y como la de qué hay que añadir exactamente al estado perceptivo interno para que se convierta en un caso de visión. Pero todas estas cuestiones y dificultades pueden evitarse en la medida en que neguemos el supuesto en el que están basadas, o sea, que hay un núcleo residual que es común a las dos situaciones mencionadas. El argumento de la ilusión trata de establecer, precisamente, la existencia de este núcleo residual compartido. Pero un filósofo de la mente que adopte el externalismo (véase 9.3) negaría la existencia de ese tipo de núcleo común, y diría que, aunque no podamos decidir si estamos viendo o alucinando, ello no quiere decir que ambas situaciones tienen un núcleo idéntico común. Se trata más bien de situaciones diferentes, a pesar de ser indistinguibles. El hecho de que, a veces, no podamos señalar sus diferencias, no quiere decir que sean «internamente» idénticas. Un externalista diría que ver es un tipo de estado mental sólo posible cuando se da un objeto externo adecuado. Que uno esté o no

en tal estado dependerá, por tanto, de si se da o no el objeto relevante. Pero el objeto no puede separarse del estado de tal modo que podamos preguntarnos si está o no causando el estado de la manera adecuada. No es posible que el objeto cause *ese* estado de un modo inadecuado. Hay otro tipo de estado que podría ser causado de esa manera inapropiada, pero eso es otra cuestión. [McDowell (1982) muestra la importancia de esta posición externalista.]

El atractivo del externalismo en filosofía de la mente, para nuestros propósitos, radica en su respuesta intransigente al argumento de la ilusión y en la resultante facilidad con la que puede manejar la cuestión, que de otro modo resulta embarazosa, sobre la relación causal que se requiere entre el estado perceptivo y el objeto. Pero este tipo de externalismo todavía no está bien elaborado. Su alternativa internalista es, con mucho, el punto de vista hoy en día más habitual.

11.5. PERCEPCIÓN, CAUSALIDAD Y JUSTIFICACIÓN

En 11.2 argumentábamos que un estado perceptivo (si hay tal cosa, a pesar de las protestas del externalismo) no es tanto un tipo complejo de sensación como una forma característica de creencia. Ahora, como epistemólogos, debemos preguntarnos por qué confiar en creencias de esa forma característica.

Hay dos cuestiones aquí que es importante distinguir. La primera es la de bajo qué condiciones estaría justificada una creencia perceptiva. La respuesta a esta cuestión es coherentista. Una creencia perceptiva está justificada si y sólo si el conjunto del que es miembro es; más coherente con ella que sin ella. Pero hay otra cuestión que comenzar mos a considerar en 8.5, al discutir la seguridad antecedente y la subsiguiente. Era la cuestión de si (y, si la tienen, por qué) las creencias perceptivas tienen cierto grado de seguridad antecedente, de si, en calidad de perceptivas, va a resultar más difícil desbancarlas. Podríamos decir que las creencias perceptivas están normalmente justificadas, o que ver cómo son las cosas a nuestro alrededor es, normalmente, saber cómo son. Si tal cosa es cierta, podría quizá seguirse que toda creencia perceptiva tiene, como tal, un buen comienzo, una oportunidad mayor de probarse, en último término, justificada. Pero, si esto es así, ¿cuál es la razón?

Podríamos comenzar sugiriendo, con Nozick, que nuestras creencias perceptivas son normalmente conocimiento porque tales creencias, normalmente, siguen el rastro de la verdad. Pero esta respuesta

tiene dos defectos. En primer lugar, la cuestión era, en realidad, la de por qué es normal que las creencias perceptivas sigan el rastro a la verdad. En segundo lugar, esta respuesta al modo de Nozick parece, como sería de esperar, una respuesta externalista, externalista en el viejo sentido del contraste externalismo/internalismo de 9.1–2. El hecho en virtud del cual mis creencias perceptivas deben contar como conocimiento (un hecho sobre su capacidad de seguir el rastro a la verdad) no es uno del que yo, como creyente, necesite ser consciente. Pero, como todavía no tenemos ninguna razón, distinta a la mera pereza, para ser externalistas, no podemos contentarnos con ninguna respuesta de este tipo.

Un segundo intento de responder nuestra cuestión adopta la perspectiva de que es una necesidad conceptual que nuestras creencias sean normalmente, o en su mayor parte, verdaderas. [Ésta es una sugerencia que hacen Hamlyn (1970), pp. 177-183, y Shoemaker (1963), pp. 229-239.] Si esto es así, cualquier apelación a ello debe ser externalista, dado que la mayoría de la gente, filósofos incluidos, no lo cree. Pero los argumentos en su favor no son concluyentes. Hay dos. El primero es una apelación a los cánones para la traducción correcta (véase 7.4). Se sugiere que nunca estaríamos en lo correcto al traducir una palabra extranjera como «ver» o «percibir» si la mayoría de las que estuviéramos considerando, al hacer tal cosa, como afirmaciones perceptivas resultaran ser (de acuerdo con nuestros propios criterios) falsas. Por tanto, sólo podemos entender como perceptivo un tipo de proceso cuyos resultados sean normalmente verdaderos. Pero este argumento es, con mucho, demasiado fuerte, y hace un uso inválido del principio de que la traducción debe preservar la verdad en la medida de lo posible. Del mismo modo, podríamos probar la imposibilidad de atribuir creencias morales a unos nativos, a menos que las creencias morales que les atribuyéramos fueran normalmente verdaderas (de acuerdo con nuestros propios criterios). Un argumento mejor es el de que los informes que los demás dan sobre sus estados perceptivos son una parte muy importante de nuestros criterios para los enunciados sobre cómo es el mundo. Por tanto, es imposible que el mundo sea, en líneas generales, distinto al modo en que la gente ve normalmente que es. Hay una dificultad general sobre el uso de este argumento aquí, dado que, cuando menos, el argumento está coloreado de tintes antirrealistas, mientras que tratamos de trabajar en el seno de un punto de vista de realismo directo. Aparte de ello, el argumento parece exagerar el papel del acuerdo. Es posible que no tengamos ninguna esperanza de construir una teoría de cómo el mundo es, por nosotros mismos, sin recurrir a la

presencia de otros. Pero no se sigue que el papel de los otros se limite a (o se defina en términos de) la necesidad del acuerdo. Y sin la necesidad del acuerdo no queda esperanza alguna para la idea de que nuestras creencias perceptivas *deben* ser (normalmente) verdaderas.

Una tercera respuesta, particularmente apropiada para el coherentista, proviene de las discusiones de 8.5. Sugerimos allí que una adhesión más empecinada a las creencias perceptivas, una que fuera más exigente a la hora de aceptar el carácter refutante de las evidencias en su contra, lleva, y se espera que lleve, a una mayor coherencia en el conjunto de creencias. Es algo similar a la justificación coherentista de los diversos principios de inferencia considerada en 8.3. Pero hay problemas para este punto de vista.

Podríamos admitir que, para alguien que sostenga esta creencia implícita, habrá un sentido en el que sus creencias perceptivas estén (normalmente, de un modo no definitivo) justificadas. Y podríamos admitir que, como asunto de hecho, todos nosotros sostenemos implícitamente esta creencia. Pero ¿qué sucedería si hubiera alguien que no la mantuviera? Lo que tratamos de mostrar, sin apartarnos del internalismo, es que todo el mundo tiene una razón internalista para considerar las creencias perceptivas justificadas en tanto que perceptivas; por tanto, la razón no puede estar en la existencia de una creencia que alguien pudiera no tener. Intentamos mostrar, de algún modo, que, incluso si las personas no consideran justificadas las creencias perceptivas, deberían hacerlo.

¿Es verdad que un aumento en la confianza en las creencias perceptivas incrementa la coherencia? Si lo es, podemos argumentar que todo el mundo debería creerlo, a causa del incremento de coherencia que resulta de la adición de esa creencia. Por tanto, dentro de los límites del internalismo y el coherentismo, hay una manera de argumentar desde la verdad a la justificación. Pero ¿qué nos dice que la proposición crucial es verdadera? Sin ninguna razón para creer que lo sea, no tenemos ninguna razón para suponer que habrá un incremento de la coherencia en el conjunto de creencias si la incluimos. Y esto vuelve el argumento del revés. En ausencia de razones para creerla verdadera, esta respuesta a la cuestión original se convierte en una hipótesis sin soporte. Es posible que pueda encontrarse alguna manera de mejorar la respuesta coherentista a nuestro problema, ello estaría de acuerdo con el programa de este libro. En caso contrario, deberemos dirigirnos a un tipo distinto de respuesta.

Una cuarta respuesta descansa sobre el elemento causal introducido en las secciones previas. Una diferencia entre creencias perceptivas

y otras es que concebimos la percepción como algún tipo de respuesta al mundo, y el mundo como algo que está frente a nosotros abierto a nuestra inspección. En un caso adecuado, la razón principal por la que vemos el mundo de un modo determinado es que el mundo es así. Una creencia perceptiva de que *p*, entonces, puede tener como su causa principal el hecho de que *p*. Esto diferencia las creencias perceptivas de las demás. No es posible percibir que todos los hombres son mortales o que $e = mc^2$; tales hechos no pueden ser la causa más importante de mi creencia, incluso aunque tengan algún tipo de papel causal.

Esto no quiere decir que todas las creencias perceptivas tengan como causas más importantes los hechos relevantes. Si tal cosa sucediera, el error perceptivo sería imposible. Lo importante es que tal cosa puede suceder en cada caso, no que suceda en todos ellos.

Esta diferencia causal entre creencias perceptivas y de otro tipo no quiere decir, por sí misma, que haya mayor probabilidad de que las creencias perceptivas sean verdaderas; la observación sobre la causalidad no es estadística. Pero podemos usar una idea nueva de vínculo directo, de un modo más cuidadoso, para decir que la percepción (cuando tiene éxito) es un vínculo particularmente directo con el mundo. Esta idea de un vínculo directo es distinta a la que hemos estado usando hasta este momento, pero todavía no es un sustituto débil de la infalibilidad. Todo lo que podemos decir es que la percepción es la forma más segura de contacto con un mundo que se nos presenta directamente.

Lo que equivale a decir que una creencia perceptiva verdadera (es decir, que tiene éxito) es normalmente conocimiento. Si hay algún conocimiento del mundo que nos rodea, la percepción es conocimiento. ¿Cómo utilizar esto para mostrar los fundamentos de la justificación de las creencias perceptivas? Hemos abandonado el intento de probar directamente que sea muy probable, o más probable, que las creencias perceptivas sean verdaderas. En vez de ello, confiamos en una creencia que todo el mundo tiene (y tiene derecho a tener) cuando está percibiendo: es decir, que su estado perceptivo, si tiene éxito, constituye una forma peculiarmente directa de contacto con el mundo.

Podríamos usar tal cosa para mostrar la justificación de la creencia perceptiva, por apelación a una sugerencia que hicimos en 3.2. Mencionamos allí un vínculo posible entre la justificación y el conocimiento, usando el esquema siguiente:

> Una creencia está justificada si y sólo si, en las circunstancias C, constituiría conocimiento.

Y propusimos el caso siguiente de ese esquema, que era el único inmediatamente obvio:

$$CJap \equiv (p \Box\!\!\rightarrow Sap).$$

Y ahora podemos ver en nuestro análisis de la percepción un caso especial de ese punto de vista. La verdadera razón por la que la justificación va unida a las creencias perceptivas es la razón por la que, si fueran verdaderas, constituirían conocimiento.

11.6. REALISMO DIRECTO Y COHERENTISMO

En esta sección final, argumento que un antirrealista puro como Dummett se compromete con el fenomenalismo en teoría de la percepción, que ya consideramos insatisfactorio (11.1); que un coherentista debe ser realista directo, antes que indirecto; y que el realismo científico directo resultará, en último término, más fuerte que su alternativa ingenua.

¿Qué se quiere decir cuando se habla de realismo perceptivo y cuál es la relación entre esa forma de realismo y el realismo metafísico, considerado en 1.4 y en 9.5? El realismo perceptivo sostiene que los objetos pueden existir y conservar (algunas de) sus propiedades no percibidas. El realismo metafísico sostiene que hay, al menos, algunas proposiciones que son verdaderas o falsas, pero cuyo valor de verdad no está determinado ni siquiera cuando se tiene en cuenta toda la posible evidencia. ¿Es posible adoptar consistentemente un realismo sin el otro?

En 9.5 distinguimos formas diversas de antirrealismo metafísico, desde el solipsismo a lo que denominamos antirrealismo puro. El solipsista, el idealista y el fenomenalista deben adoptar claramente un antirrealismo perceptivo. ¿Qué sucede con el antirrealismo puro? Podría parecer que puede ser realista perceptivo, en la medida en que su realismo sea directo más que indirecto, dado que el realismo perceptivo es el punto de vista de que la posibilidad de la experiencia perceptiva tiene un fundamento que es independiente y distinto de esa posibilidad, o sea, la existencia continuada del mundo material. Si podemos concebir qué sería la existencia de tal fundamento, sin poder concebirla como algo irreconocible, habremos unido el realismo perceptivo con el antirrealismo puro. Pero, en este caso, nuestro realismo perceptivo deberá ser directo, porque un realista indirecto encontrará difícil evitar decir que el mundo material, que es el objeto de percep-

ción, es algo independiente de la experiencia posible, y, por tanto, puede darse al margen de ella.

Sin embargo, creo que hay una tensión incómoda en esta amalgama de realismo perceptivo y antirrealismo puro y que, en último término, es letal. El antirrealista puro más consistente no concederá que podamos concebir un estado de cosas que sea, a veces, reconocible y, otras veces, irreconocible. Si queremos ser consistentes deberemos decir que no podemos concebir en qué pueda consistir que algo sea irreconociblemente verdadero en *alguna* circunstancia, porque nuestra comprensión de la verdad de esa oración está vinculada de una forma tan estrecha a lo que consideramos como evidencia para ella, que no podemos separar ambas y suponer que la verdad está presente sin que lo esté la evidencia (o viceversa). Pero, ciertamente, el realista directo, al suponer que existe un fundamento *independiente* y distinto de la posibilidad de la experiencia, está suponiendo que ese fundamento podría estar presente de un modo que fuera irreconocible, o que podría parecer presente cuando en realidad estuviera ausente. El concepto de independencia incluye, en este caso, el de variabilidad independiente. Lo que es inconsistente con el antirrealismo puro, que, por tanto, se ha de comprometer con una posición en filosofía de la mente que ya hemos rechazado, o sea, el fenomenalismo.

Argumenté en 9.5 que el coherentista todavía podría ser un realista de la forma más fuerte; y es probable que tal coherentista sea también realista en teoría de la percepción. ¿Qué tipo de realista debería ser? Mi opinión sería que el realismo directo es perfectamente acorde con el espíritu del coherentismo, mientras que el realismo indirecto vuelve a abrir una brecha inaceptable entre verdad y justificación.

La cuestión central es el sentido que demos a la idea del coherentista de que justificación y verdad no son propiedades de tipos radicalmente diferentes; o que la justificación, a medida que se incrementa, se aproxima a la verdad (aunque nunca puede llegar a ella). El realista que piense que las propiedades de los objetos físicos son, a la vez, directamente observables y capaces de perdurar cuando nadie las observa acepta que el mismo estado de cosas que observa puede darse aunque nadie lo observe. Esto es realismo directo. Sin embargo, el realista indirecto sostiene que no podemos concebir el objeto que aprehendemos directamente como algo que pueda continuar existiendo cuando nadie lo percibe. El mundo que perdura, cuya naturaleza es la que convierte en *verdaderos* nuestros estados de aprehensión, es, por tanto, distinto a cualquier concepción del mismo por apelación a la cual se *justifiquen* nuestras creencias perceptivas. Es de presumir

que la verdad sea una relación entre el objeto directo y el indirecto, mientras que la justificación será una relación entre los objetos directos de aprehensión; una creencia perceptiva estará justificada en la medida en que se ajuste bien con otras creencias. Pero esto crea una distinción entre verdad y justificación que, desde el punto de vista del coherentista, no puede ser correcta.

Históricamente hablando, el realismo indirecto se ha sentido atraído por el tipo de escepticismo que surge cuando pensamos que, por todo lo que sabemos, detrás del mundo de apariencias puede yacer un mundo real completamente diferente, o incluso ningún mundo real en absoluto. Los realistas indirectos han tendido a aceptar esa posibilidad y han argumentado que, a pesar de ello, podemos conocer el mundo real gracias a nuestra aprehensión del mundo de apariencias. Por el contrario, para él realista directo, la distinción entre el mundo tal y como es y nuestra manera de percibirlo, aunque se acepte, es mucho menos radical. Y esto es así porque la naturaleza del mundo real es exactamente del mismo tipo que la naturaleza de las propiedades que percibimos que tiene. Las propiedades que, gracias a la percepción, le atribuimos al mundo son propiedades que el mundo puede tener y que pueden perdurar cuando nadie lo percibe.

Así que, por esas dos razones, creo que el coherentista sólo puede ser consistente si adopta una forma de realismo directo en teoría de la percepción. Lo cual, por supuesto, no es lastre alguno, porque el realismo directo es, sobre razones independientes, la teoría más verosímil.

¿Cuál de las formas de realismo directo deberá preferir? En 10.3 di argumentos que apuntaban que el realismo directo ingenuo era más consistente que su alternativa científica. Pero esos argumentos pueden parecer ahora menos contundentes. Eran dos: en primer lugar, la duda sobre la distinción entre cualidades primarias y secundarias, y, en segundo término, el temor a que el uso de esa distinción provocara que el realista directo científico cayera en el realismo indirecto para explicar la percepción del color. Pero la discusión del error perceptivo en 11.3 parece haber alejado tal temor. Con lo que hemos vuelto a abrir la posibilidad de que el coherentista acepte la distinción entre cualidades primarias y secundarias y opte por el realismo científico directo.

LECTURAS COMPLEMENTARIAS

Grice (1961), sobre los elementos causales en la teoría de la percepción y su relación con el escepticismo.

Cornman (1975, cap. 7) examina el realismo indirecto (clasificando a Sellars como realista indirecto) y lo rechaza.

Jackson (1977, cap. 6) responde con ciertas objeciones al realismo indirecto.

McDowell (1982) argumenta en favor de una respuesta externalista al argumento de la ilusión.

Snowdon (1981) usa un punto de vista externalista para cuestionar la necesidad de una teoría de la percepción completamente causal. Se alude a los efectos del externalismo en epistemología en el comienzo de Sellars (1973).

Strawson (1974, cap. 4) continúa el trabajo de Grice sobre la teoría causal.

Weiskrantz (1980) proporciona un análisis introductorio del fenómeno de la visión ciega.

Rock (1984) es una introducción clara a la psicología de la percepción, accesible a los filósofos.

Respecto a la elección entre una teoría de la percepción como «pura creencia» y una «mixta», véase Peacocke (1983, cap. 1). Sellars (1963, pp. 155-156) protesta contra la asimilación de las sensaciones a pensamientos.

12. MEMORIA

12.1. TEORÍAS DE LA MEMORIA

Podríamos pensar que percibir es aprehender la naturaleza del medio en el que estamos ahora, y que, de un modo semejante, recordar es aprehender la naturaleza de nuestro pasado. Las diferencias y semejanzas entre memoria y percepción serán un tema constante en este capítulo; las semejanzas parecerán mucho mayores que las diferencias. Pero ya sabemos que el contraste establecido en la primera oración está mal concebido; no es necesario que el medio circundante que aprehendemos en la percepción esté temporalmente presente, y encontraremos razones para dudar de que la memoria deba restringirse a la aprehensión de nuestro pasado.

El alcance de la semejanza entre la memoria y la percepción está testificado por la existencia de doctrinas análogas, en la teoría de la memoria, para cada una de las tres familias de teorías de la percepción. Podemos distinguir, de nuevo, entre realismo directo, realismo indirecto y fenomenalismo. Por sí mismo, este hecho ya constituye una tentación para concebir la memoria como la percepción del pasado. No es completamente necesario que nos resistamos a ella, por más que también puedan existir formas de memoria que no puedan concebirse como perceptivas.

Al considerar las pretensiones de las tres teorías, nos encontraremos a menudo con repeticiones de los argumentos que se formularon en teoría de la percepción. Lo que quiere decir que podemos ir un poco más deprisa. Por otra parte, con ello tendremos un contexto diferente en el que volver a evaluar la fuerza de tales argumentos.

Como paso previo, necesitamos aclarar qué se quiere decir cuando se habla de realismo en esta nueva área. Nuestro análisis debe y puede ajustarse al que se dio en 10.6 (que era una mejora de la versión inicial de 10.2). El realismo sobre la memoria es el punto de vista de que tanto la presencia de ciertos recuerdos, como la accesibilidad o la posibilidad de recuerdos distintos a los que de hecho tenemos, están fundamentados en algo completamente diferente a los recuerdos mismos,

algo cuya naturaleza explica la posibilidad de éstos. Ese algo es el curso del pasado, la historia previa del universo. En el recuerdo conservamos o recobramos la aprehensión de algo que es independiente de nuestra aprehensión, dado que las propiedades del pasado no dependen de nuestra capacidad actual de recordarlo.

12.2. REALISMO INDIRECTO

El realista indirecto sostiene que recordar es aprehender indirectamente el pasado. Cuando recordamos, hay un objeto directo de aprehensión que funciona a modo de intermediario; es la imagen de la memoria. La imagen de la memoria es nuestro objeto directo interno.

Los argumentos a favor del realismo indirecto sobre la memoria tienden a parecerse a los argumentos en favor del realismo indirecto en la percepción. En ambos casos se da el impulso a transformar el contenido de la aprehensión en un segundo objeto; se apela al argumento del intervalo temporal, que es especialmente tentador en teoría de la memoria, y se utiliza una versión del argumento de la ilusión. Ya hemos visto que estos argumentos no son concluyentes, y la atracción adicional del argumento del intervalo temporal no los convierte en concluyentes. No hay ningún argumento que nos obligue a aceptar la existencia de un objeto directo e intermediario del recuerdo.

Este rechazo del realismo indirecto es demasiado precipitado, especialmente si recordamos que los realistas indirectos han sido abrumadora mayoría en la reciente historia de la filosofía. Pero podemos completarlo con algunas observaciones.

En primer lugar, los verdaderos problemas para el realista indirecto en teoría de la percepción surgían cuando intentaba proporcionar un análisis de la doble aprehensión y del doble objeto. Esos problemas se plantean, exactamente del mismo modo, en el caso de la memoria. Si la imagen es el objeto directo intermedio, ¿cómo se las arregla para funcionar como tal?

En segundo lugar, la intermediación entre la mente que recuerda y el suceso pasado es doble. Ya que los argumentos en favor del realismo indirecto son los mismos en el caso de la memoria y en el caso de la percepción, es probable que el realista indirecto encuentre dos objetos de aprehensión entre la mente que recuerda y el objeto original. Está, en primer lugar, el objeto original directo, intermediario entre el perceptor en ese momento y la cosa percibida. Además, hay un segundo objeto directo que se le presenta ahora a la mente que

recuerda y que, de algún modo, está relacionado (probablemente se parece) con el anterior objeto directo que ahora está más allá de las posibilidades de aprehensión, dado que ha dejado de existir (por el argumento del intervalo temporal). Hay dos razones por las que no podemos suponer que el objeto directo anterior vuelve a aparecer tras un lapso en el que ha permanecido oculto. En primer lugar, no es obvio que cosas del tipo de las imágenes de la memoria puedan retornar a la existencia tras un lapso de tiempo. Más bien deberíamos suponer que aparece una imagen distinta que se asemeja a la anterior, de otro modo estaríamos tentados a decir que la imagen recurrente ha estado «almacenada» de un modo u otro durante los intervalos entre nuestra aprehensión de ella. Con ello estaríamos concibiendo la memoria como suerte de almacén de imágenes no usadas, lo que parece una concepción muy poco atractiva. En segundo lugar, incluso si una imagen puede, en principio, volver a la existencia después de un intervalo, es normal que las imágenes de la memoria difieran tanto de las originales que (bajo el supuesto de que éstas son los objetos directos originales) sea difícil considerarlas como supervivientes de ellas. Tenemos, pues, que tanto el intervalo temporal como las diferencias cuentan contra la identidad de las diversas imágenes a través del tiempo. Pero la doble intermediación entre la mente y el objeto indirecto acentúa todos los problemas sobre la relación entre objeto directo e indirecto en el caso de la percepción. Por ejemplo, considerábamos en 11.2 la objeción de que la existencia de un objeto de aprehensión intermediario entre nosotros y el mundo hará imposible que sepamos cómo es el mundo. Cualquiera que esté impresionado por esa objeción la encontrará incluso más contundente en el caso de la memoria.

La tercera cuestión que se plantea habitualmente como objeción al realismo indirecto está relacionada con nuestra capacidad de distinguir entre la memoria y la imaginación. El realismo indirecto concibe ambas facultades como productoras de imágenes. Pero consideramos que algunas de estas imágenes son imágenes–de–memoria, mientras que las otras son meros productos de la imaginación. ¿Qué criterios utilizamos para distinguirlas?

La cuestión se plantea a menudo como demanda de un criterio infalible. La idea es la de que, sin un criterio infalible que distinga entre memoria e imaginación, seremos incapaces, en un caso determinado, de decir si estamos recordando esas cosas o sólo imaginándolas. Y, si no podemos decir eso, tampoco podemos decir que alguien que recuerde el pasado tiene conocimiento del pasado. ¿Cómo es posible

que la memoria nos dé conocimiento del pasado, si ni siquiera puede discriminar con seguridad si estamos recordando o sólo imaginando?

La preocupación sobre estos retos escépticos es real, pero está mal dirigida. Los peligros de este tipo de argumento escéptico del error son reales, pero no debemos suponer que la manera de eliminarlos sea el descubrimiento de un criterio infalible. No hay criterios infalibles, y ninguna teoría debería ser rechazada por su incapacidad de proporcionar uno. Todavía necesitamos enfrentarnos al argumento escéptico del error, pero debemos enfrentarnos a él en términos generales, no sólo con relación a este tema específico, y no debemos confundir el intento de vérnoslas con tal argumento con la búsqueda de un criterio infalible.

Nuestra cuestión, entonces, es la de cómo distinguir de hecho entre las imágenes que son productos de la memoria y las otras. Hay dos respuestas clásicas a esta cuestión. La primera se la debemos a Hume (edición de 1967, libro 1, 1.3). Sugirió que las imágenes de la memoria son, en general, más vividas y más fuertes que las demás. No está del todo claro lo que Hume quería decir con esto. Después de todo, hay un sentido en el que los productos de la imaginación son más vividos, en general, que las imágenes que son restos difusos del pasado. Es posible que Hume no lo negara. Una manera más provechosa de entenderle sería recordar que él también utilizaba el mismo criterio para distinguir entre la creencia y la imaginación. Los contenidos de nuestra creencia son más vividos y fuertes que los de la imaginación. Y lo que con ello quería decir es que nuestras creencias influyen en nuestras acciones de un modo en que no lo hacen los contenidos de nuestra imaginación. Si esto es correcto, lo que Hume sugería es que la memoria es una forma de creencia. Nuestra capacidad de distinguir la memoria de la imaginación es un caso de la capacidad más general de distinguir lo que creemos de lo que no.

Antes de pasar a evaluar esta respuesta, consideremos una alternativa. Russell sugirió (1921, cap. 9) que nuestro criterio en este punto es un sentimiento de familiaridad que acompaña normalmente a las imágenes de la memoria y no a las de la imaginación. El sentimiento puede ser muy fuerte (e incluso puede estar mal dirigido, como en el sentimiento de *déjà vu*) o muy débil. Pero está presente normalmente y constituye el criterio que estamos buscando.

El problema con la respuesta de Russell no es tanto el que la presencia o ausencia de este sentimiento no sea una marca completamente fiable (lo que sería de esperar, dado que no hay marcas completamente fiables). El problema, más bien, es el de distinguir el sentimien-

to de familiaridad y la creencia de que la imagen presente representa, de algún modo, el pasado. Russell pretende que el sentimiento y la creencia son distintos (ibid., p. 169), estando la creencia basada siempre en la presencia del sentimiento. La réplica es que, en ese caso, no hay un sentimiento de familiaridad distinto a la creencia que se supone que ese sentimiento fundamenta. El denominado sentimiento *es* exactamente la creencia que se supone que fundamenta.

Con ello parece que la respuesta de Russell no es tal respuesta. Si la cuestión era qué criterio usamos para distinguir entre memoria e imaginación, la respuesta no puede ser que en un caso creemos y en el otro no. La creencia no puede ser el criterio, dado que es exactamente el criterio de la creencia lo que estamos buscando. La creencia, por tanto, parece que no tiene criterio.

La misma conclusión podemos extraer del ejemplo de Hume. Si estamos de acuerdo en que el recuerdo es una forma de creencia que va unida a una imagen presente, no es posible encontrar un criterio por el cual podamos adscribir creencia en un caso y no en el otro. La adscripción de creencia carece, de nuevo, de criterios.

Parece, por tanto, que la cuestión de qué imágenes son de memoria no puede responderse; pero es una cuestión respecto a la que el realista indirecto tiene razones para esperar una respuesta. Veremos que el realista directo sí tiene respuesta a la cuestión, aunque la razón sea que la cuestión ya será otra. La cuestión del realista directo no trata de imágenes, sino de creencias. ¿Qué creencias son creencias-de-memoria?

El obstáculo principal para aceptar el realismo indirecto en teoría de la memoria es que parece que tenemos dos alternativas: o bien adoptar el realismo indirecto, tanto para la percepción como para la memoria, o bien adoptar el realismo directo en ambos casos. Una teoría mixta, directa en un caso e indirecta en otro, no será atractiva porque los argumentos aceptados en uno de los casos se extenderían irremediablemente al otro. Habiendo argumentado a favor del realismo directo en teoría de la percepción, no es probable que seamos consistentes si aceptamos el realismo indirecto para la memoria.

12.3. REALISMO DIRECTO

El realista directo sostiene que en el caso de memoria, como en el de la percepción, nuestra aprehensión del pasado es directa. No hay ningún objeto intermediario «interno», en virtud de cuya aprehensión podamos aprehender otras cosas de un modo indirecto. Las imágenes–

de–la–memoria, en el caso de que se den alguna vez, no funcionan de la manera en que sugiere el realista indirecto. Las imágenes no son tan importantes para el recuerdo como los filósofos han tendido a suponer, e, incluso, cuando sí hay una imagen, ella misma no es un objeto de aprehensión, sino (parte de) la manera de que aprehendemos el suceso pasado.

De modo que el realista directo reduce el papel de la imagen u objeto interno, tanto en la memoria como en la percepción. ¿Es razonable? ¿No se arriesga a perder de vista lo realmente característico de la memoria?

Para responder a esta pregunta es necesario que investiguemos lo que se han considerado como formas diferentes de recuerdo, para ver cómo están relacionadas y qué papel desempeñan las imágenes en cada una de ellas.

Una forma de memoria, de la que hasta ahora se ha dicho bien poco, pero que es la más común, es la *memoria fáctica*. La memoria fáctica es conocimiento fáctico, conocimiento de que *p*, que se ha adquirido en el pasado; este conocimiento puede haberse perdido (olvidado) y recuperado, o puede no haberse perdido nunca y haber estado simplemente conservado. Parece que mucho de nuestro conocimiento está constituido por memoria fáctica. Pero no todo; nuestro conocimiento de lo que estamos percibiendo en este preciso momento no es memoria fáctica, incluso aunque nos fuera imposible tener tal conocimiento si no tuviéramos esa memoria.

No es preciso que los hechos con los que está relacionada la memoria fáctica sean relativos al pasado. Puedo recordar que estaré en Londres el próximo fin de semana, o que hay más de una variedad de pájaro carpintero. Hay, pues, un sentido importante en el que la memoria no está limitada al pasado y mucho menos al pasado propio.

La memoria fáctica no depende, en modo alguno, de que se den las imágenes adecuadas o los objetos directos «internos». Podemos advertir que sucede así porque gran parte de nuestra memoria fáctica es inimaginable en un sentido especial: es relativa a hechos de los que no caben imágenes relevantes. Por ejemplo, el conocimiento matemático está constituido, en gran parte, por memoria fáctica. No es posible tener una imagen del hecho de que 2 + 2 sea equivalente a 4.

Parece, pues, que la memoria fáctica, a pesar de su carácter prevalente, no es lo que concibe el realista indirecto. No es particularmente relativa al pasado, y las imágenes no son, en ella, ni imprescindibles ni especialmente relevantes. Seguramente ello demuestra que hay otra forma de memoria sobre la que sí hablaba el realista indirecto. Sea

como sea, parece obvio que es así. La memoria fáctica es relativa a la conservación o la recuperación de algún tipo de conocimiento. Si esto fuera todo lo que es la memoria, no tendríamos demasiadas esperanzas de extraer paralelismos interesantes entre memoria y percepción. Lo que es más, no hay nada en la noción de memoria fáctica que pueda dotar de sentido a la idea de que sólo puedo recordar algunos de los sucesos u objetos del pasado. Y es cierto que sí creemos que, en algún sentido, nuestra memoria está limitada a nuestra propia historia pasada. Lo que no es decir que sólo podemos recordar lo que conocimos anteriormente, sino que sólo podemos recordar sucesos y objetos que aprehendimos alguna vez anteriormente. Si hay un sentido en el que esto es así, queda espacio para las analogías entre percepción y memoria y para las discusiones entre realismo directo e indirecto sobre, por ejemplo, el papel y la función de la imaginación.

Una manera no muy buena de prestar atención a la distinción entre memoria fáctica y lo que podríamos denominar memoria *perceptiva* sería apelar al lenguaje. ¿No nos ofrece el lenguaje la distinción entre que te acuerdes de que hiciste algo y que te acuerdes de ti haciendo algo? ¿No podemos suponer que para el segundo caso, y no para el primero, es necesario que puedas recordar el suceso «con la misma conciencia de él que tuviste en la primera ocasión»? De un modo semejante, podríamos distinguir entre recordar que algo parecía similar a algo y recordar cómo parecía, o recordar que el partido terminó empatado y recordar el partido que terminó empatado. Pero las distinciones lingüísticas no se adecuan completamente a la distinción entre memoria fáctica y perceptiva, aunque sólo sea porque la memoria perceptiva puede ser descrita perfectamente en términos de recordar-que.

Debemos ser cuidadosos y no dar demasiado por supuesto en la relación entre memoria fáctica y perceptiva. Por todo lo que sabemos hasta ahora, la memoria perceptiva podría ser sólo un tipo de memoria fáctica. Supongamos las siguientes definiciones provisionales:

> *a* recuerda fácticamente que p si y sólo si el conocimiento actual de *a* de que *p* se fundamenta en el conocimiento previo de *a* de que *p*.
>
> *a* recuerda perceptivamente que *p* si y sólo si el conocimiento actual de *a* de que *p* se fundamenta en la percepción previa de *a* de que *p*.

Si añadimos a ellas la idea de que la percepción misma puede ser una forma de conocimiento (véase 11.5), nos encontramos con el

resultado de que la memoria perceptiva es un tipo especial de memoria fáctica. Todavía habría una distinción entre ellas, pero no un contraste. [Martin y Deutscher (1966) ofrecen razones para pensar que las definiciones anteriores son demasiado restrictivas.]

Es preciso que consideremos ahora lo que tenga que decir el realista directo sobre la memoria perceptiva, la capacidad de ver de nuevo, «en el ojo de la mente», lo que vimos antes, volver a experimentar los propios sentimientos o volver a vivir las propias acciones. Esta capacidad está fundamentada en las propias experiencias anteriores (percepciones) y, por tanto, cualquier análisis de ella va a depender del análisis previo de la percepción. Ya hemos visto el análisis del realista indirecto: la memoria perceptiva es la aprehensión de una imagen presente que se parece a una imagen previa que representaba un objeto. No es probable que el realista directo, que ha rechazado el intermediario anterior, admita este último. ¿Qué dificultades adicionales hay en esta posición?

Puede parecer que hay un problema en la posibilidad de que nuestra aprehensión se altere a lo largo del tiempo, incluso para mejorar. Parece que es posible que alguien decida, apelando a sus recuerdos, que un colega estaba presente en cierta conferencia, incluso aunque haya un sentido en el que no pueda decirse que, entonces, fue consciente de su presencia. La cuestión es por qué este fenómeno ha de concebirse como una objeción al realismo directo. ¿Por qué ha de comprometerse el realista directo con la tesis de que nuestra aprehensión del objeto no puede cambiar a lo largo del tiempo? O, si no, ¿por qué debe comprometerse con la tesis de que sólo puede deteriorarse? La posibilidad de que cambie a mejor debe ser un argumento a favor, y no en contra, del realismo directo, dado que supone que nuestra aprehensión actual del objeto no está limitada por la naturaleza de alguna experiencia o imagen previa, sino que puede ser mejor. ¿Cómo sería posible tal cosa si no aprehendiéramos directamente el objeto, incluso a través del tiempo?

Si el realismo directo puede sobreponerse a este tipo de objeciones, ¿cómo se verá afectado por la distinción entre teorías (de la memoria, no de la percepción) sensoriales, «mixtas» o de «pura creencia»? En el caso de la memoria fáctica, la perspectiva de la creencia pura parece ser obviamente correcta. Parece que se dan analogías entre la percepción y la memoria perceptiva, la capacidad de volver a vivir el propio pasado. ¿Deberemos adoptar, por tanto, una teoría mixta para la memoria perceptiva y mantener intacta la analogía?

Muchos filósofos se resisten menos a la teoría de la creencia pura

de la memoria en general que a las teorías de la creencia pura de la percepción. La razón es que en la percepción la *apariencia* es central, mientras que en la memoria no es necesario que haya un modo de aparición o de reaparición del pasado a quien está recordando. Lo que quiere decir, como se admite, que en la memoria fáctica las apariencias son irrelevantes. Todavía queda el fenómeno obstinado de la memoria perceptiva, en el que las apariencias *sí* son importantes. Y, dada la naturaleza distintiva de la memoria perceptiva, el mejor camino parece ser el de preservar la analogía entre percepción y memoria, adoptando una teoría mixta en ambas áreas. Debemos suponer que la memoria perceptiva es una forma distintiva de creencia sobre el propio pasado. Según este análisis, la memoria perceptiva, en tanto que creencia, es una forma de memoria fáctica; pero se trata de una forma distintiva. Ambas son creencia, pero, aunque alguien podría tener memoria fáctica sin memoria perceptiva, nada podría disfrutar de memoria perceptiva sin la memoria fáctica.

12.4. FENOMENALISMO

Podemos distinguir, como antes, entre fenomenalismo eliminativo y reductivo. El fenomenalismo eliminativo sostiene que no existe tal cosa como el pasado; no hay nada más que el hecho de que sucedan en el presente experiencias de cierto tipo. El fenomenalismo reductivo mantiene que *sí* existe el pasado, pero que no es nada más que un complejo de ciertas experiencias presentes. Para que un suceso se dé en el pasado es necesario que tengamos o seamos capaces de tener una experiencia–de–recuerdo en el presente; el pasado consiste en el hecho de que las experiencias–de–recuerdo nos sean accesibles. [Para este punto de vista, véase Ayer (1946), pp. 101-102.]

Este tipo de teoría sufre de los mismos defectos que los de su contrapartida en teoría de la percepción. ¿Qué explica que se dé una experiencia–de–recuerdo? Normalmente, responderíamos esta cuestión apelando a un suceso previo percibido, que se recuerda ahora. Ésta es la explicación que no se puede permitir el fenomenalista, que se limitaría a explicar el hecho de que se dé una experiencia–de–recuerdo apelando a la accesibilidad de esa experiencia. Lo que, en último término, no parece ser ningún tipo de explicación. El fenomenalista ni siquiera puede apelar a las regularidades previas en la sucesión de experiencias–de–recuerdo. Y no porque tales regularidades estén en el pasado: hay un análisis fenomenalista de lo que está en el pasado en

términos de accesibilidad en el presente de experiencias–de–recuerdo. La apelación a las regularidades previas es, pues, consistente con el fenomenalismo, pero no logra su objetivo. No lo hace porque, como en el caso de la percepción, las regularidades previas no proporcionan el tipo de explicación que buscamos. Nos confirman que la experiencia–de–recuerdo es predecible, pero no nos dicen por qué la experiencia sucedió de un modo y no de otro.

El fenomenalismo en teoría de la memoria es, por tanto, defectuoso. Pero hay algo que podemos aprender del contraste entre fenomenalismo y realismo en este punto. Para el fenomenalismo, la memoria fáctica sólo puede existir si hay memoria perceptiva. La memoria fáctica requiere la existencia de conocimiento pasado; pero el fenomenalismo considera que el hecho de que hubiera un momento en el que yo supe que *p* consiste en la accesibilidad de experiencias de haber sabido que *p*; lo que está involucrado en la memoria perceptiva es que se dan tales experiencias. Sin embargo, para el realista parece que lo contrario sea la verdad. La memoria perceptiva no es posible sin memoria fáctica. Ello es así porque el realista considera la memoria perceptiva como una especie de memoria fáctica. Dada esta relación, podemos decir que, si no hubiera tal cosa como el conocimiento presente basado en el conocimiento pasado (memoria fáctica), no podría existir nada como el conocimiento presente basado en la percepción pasada (memoria perceptiva). El carácter estricto de este vínculo se revela en el reconocimiento de que, en casos favorables, percibir que *p* es saber que *p*.

La importancia de este contraste reaparecerá cuando tratemos de considerar las cuestiones escépticas sobre el conocimiento y la memoria. ¿Cómo puede una creencia–de–memoria ser conocimiento?

12.5. LA HIPÓTESIS DE RUSSELL

Las dificultades escépticas se plantean mejor, quizá, por medio de la hipótesis de Russell de que el mundo ha sido creado hace exactamente cinco minutos del modo exacto en que es, con una población que «recuerda» un pasado completamente irreal [Russell (1921), pp. 159-160]. Como afirma Russell, nada de lo que sucede ahora, o de lo que pudiera suceder, puede probar que el mundo no se creó de ese modo. Más importante, parece que nada de lo que está sucediendo o pudiera suceder podría contar como evidencia de que el mundo no se creó de ese modo. Por tanto, si no podría haber evidencia

contra la hipótesis de Russell, es seguro que no podemos saber que la hipótesis es falsa. Y, en tal caso, ¿cómo podemos llegar a conocer algo sobre el pasado? Si pudiéramos saber algo sobre el pasado, por ejemplo que ha habido dos guerras mundiales en este siglo, sabríamos que el mundo no ha sido creado hace cinco minutos. No sabemos esto último y, por tanto, no sabemos lo anterior.

Hay que hacer dos comentarios sobre este argumento, antes de considerar tres posibles respuestas a él. En primer lugar, el argumento es un mero caso del argumento del error (1.2), por el que antes mostrábamos que no sabemos que no somos cerebros en una cubeta. El núcleo del argumento era la existencia de dos hipótesis que nos son completamente indistinguibles; no podemos pretender saber cuál es verdad, ni siquiera tener justificación alguna para cualquier creencia al respecto. El caso que nos ocupa es exactamente así. No hay ninguna posibilidad de descubrir evidencia alguna en favor de la hipótesis de Russell, más bien que en favor de la concepción tradicional, o viceversa. En segundo lugar, debe estar claro que el primer objetivo del argumento escéptico no es la memoria perceptiva, sino la fáctica. Si el argumento es correcto, independientemente de si hay o no cosas tales como las experiencias–de–recuerdo, no habrá caso alguno de conocimiento presente que esté fundamentado en el conocimiento pasado.

La primera respuesta posible al argumento escéptico es la de Nozick. Se articula bajo el modelo de su respuesta, considerada en el capítulo 3, a las dificultades escépticas con los cerebros en cubetas. Nozick sería amigo de solucionarlas diciendo que el principio de inferencia PC^S sobre el que las que descansan es falso, por lo que podemos llegar a saber más hechos normales sobre el pasado, incluso aunque no podamos saber que la hipótesis de Russell sea falsa. Sin embargo, sugerimos en el capítulo 3 que, incluso si PC^S es falso, el movimiento de Nozick no puede alcanzar su objetivo. Hay otra manera de argumentar, por medio del principio de universalizabilidad, que, si no sabemos que la hipótesis de Russell es falsa, no seremos capaces de conocer hechos más mundanos sobre el pasado. Y, ciertamente, esto se ajusta a nuestras intuiciones. A Nozick le gustaría que aceptáramos que yo puedo saber lo que desayuné esta mañana sin poder saber si el mundo existe desde hace más de cinco minutos. Sin embargo, tenemos la fuerte intuición de que no deberíamos poder decir tal cosa.

El fenomenalismo, si fuera correcto, proporcionaría una respuesta mucho mejor. Sostiene que la existencia del pasado no es otra cosa

que la accesibilidad de las experiencias–de–recuerdo. Si esto fuera así, no podría haber una situación en la que fueran accesibles las experiencias–de–recuerdo sin que se hubiera dado ninguno de los sucesos pasados que esas experiencias parecen representar. Podemos expresarlo, como en 1.5, diciendo que no es como si hubiera dos alternativas, una en la que habría experiencias–de–recuerdo, pero sin pasado, y otra en la que habría no sólo las experiencias–de–recuerdo, sino también un pasado fundamentando esas experiencias. No podemos decir que hay una diferencia que no podemos distinguir y que no establecería diferencia alguna para nosotros. Más bien, el hecho de que no pudiéramos distinguirla y el hecho de que no establecería ninguna diferencia para nosotros muestran que no hay ninguna diferencia que distinguir. Lo único que determina que esos sucesos hayan sucedido es la accesibilidad de las experiencias–de–pasado. En tal caso, la hipótesis de Russell no proporciona ninguna razón para que aceptemos una conclusión escéptica. Es éste un ejemplo de la relativa facilidad con la que un antirrealista puede rechazar los argumentos escépticos.

De modo que el pasado, tal como lo concibe el fenomenalismo, es algo con lo que estamos en contacto directo en las experiencias–de–recuerdo. Y no hay sucesos pasados que estén fuera del alcance de la memoria. No hay hechos relativos al pasado distintos de los que determinan la accesibilidad de las experiencias–de–recuerdo. Es posible que haya oraciones sobre el pasado que pretendan expresar hechos sobre el pasado que están más allá del alcance de la memoria, pero tales oraciones no son verdaderas ni falsas. Sólo serían verdaderas o falsas si fueran accesibles ciertas experiencias–de–recuerdo de los sucesos que las oraciones describen como habiendo sucedido o no sucedido. Al no ser accesibles ese tipo de experiencias, las oraciones implicadas no son ni verdaderas ni falsas.

Esta concepción del pasado se opone radicalmente a la concepción l realista. El realista considera que ciertas maneras de existir el mundo en el pasado pueden estar más allá de los recuerdos que nos son accesibles. El pasado es más rico que las posibilidades de la experiencia–de–recuerdo. Un realista puede, aunque quizá no le sea necesario, expresar su posición diciendo que toda oración sobre el pasado es, de hecho, o bien verdadera o bien falsa, independientemente de si son o no accesibles ciertas experiencias–de–recuerdo relevantes. Ésta es una manera de defender que la Ley de Tercio Excluso, que mantiene que cualquier proposición es o bien verdadera o bien falsa, es válida para las oraciones sobre el pasado. El fenomenalista parece en la obligación de negar la Ley de Tercio Excluso, y decir que algunas oraciones

describen sucesos que ni sucedieron ni dejaron de suceder. De modo que el pasado, concebido de una manera realista, se extiende más allá de las fronteras a las que le gustaría limitarse el fenomenalista.

Una vez adoptamos el punto de vista realista de que son dos cosas completamente distintas el que haya ocurrido un determinado suceso y el que sean accesibles experiencias–de–recuerdo del mismo (concebidas como imágenes o creencias de cierto tipo), nos enfrentamos al argumento escéptico habitual. El argumento del error arranca de la observación de que, con independencia de la consistencia de nuestras experiencias-de-recuerdo, siempre existe la posibilidad de que sean falsas. Hasta ahora no hemos visto ninguna respuesta satisfactoria a este tipo de argumento. Su ataque a nuestra pretensión de conocer el pasado era de esperar, y debe ser controlado, si fuera posible, por medio de alguna estrategia epistemológica más general y no por un movimiento especial en teoría de la memoria. El fenomenalismo tiene el mérito de proporcionarnos ese tipo de estrategia.

Otra estrategia general de ese tipo se consideró en 1.4. Era el argumento «trascendental». Enfrentados a la tendencia escéptica de la hipótesis de Russell, intentamos, en primer lugar, maximizar sus efectos. Decimos, por ejemplo, que, si la hipótesis de Russell fuera verdadera, la memoria fáctica sería imposible. Y no sólo no habría posibilidad de conservar el conocimiento del que disfrutamos en el pasado, sino que tampoco habría posibilidad alguna de obtener ahora un conocimiento nuevo. ¿Cómo podría saber que eso es un fuego, si no puedo recordar lo que es un fuego? Este último movimiento, tal y como está formulado, es de dudosa validez. Pero pasémoslo por alto; lo esencial es que tratemos de maximizar los efectos del movimiento escéptico por todos los medios a nuestro alcance y que, después, señalemos que, una vez perdido todo eso, nos será muy difícil considerarnos a nosotros mismos como seres racionales. ¿Cómo podríamos considerar algo como una razón para otra cosa si no tuviéramos el tipo de información que nos proporciona la memoria fáctica? De modo que los efectos del argumento escéptico son gigantescos. Tan gigantescos que el argumento debería ser rechazado.

No he formulado este argumento exactamente del mismo modo que en 1.4. Pero la respuesta más importante a él es la misma. Las consideraciones escépticas no se debilitan porque reconozcamos el alcance de sus consecuencias. Es posible que nos sintamos más inclinados a dudar de la validez de esas consideraciones, pero la mera constatación de la enormidad de sus consecuencias no es un argumento independiente en favor de que tales consideraciones contienen

defectos importantes. No podemos eludir la tarea de diagnosticar el error del argumento escéptico, tal y como tratan de hacer los argumentos trascendentales. Por tanto, nuestras tres respuestas al argumento clásico sobre la memoria fáctica han fallado. ¿Dónde nos deja todo esto? Necesitamos una manera de tratar con el argumento del error, y todavía no hemos encontrado ninguna. Realizaré mis propias sugerencias en 15.5.

12.6. MEMORIA PERCEPTIVA Y JUSTIFICACIÓN

Como en 11.5, hay dos cuestiones que nos podemos preguntar sobre la memoria perceptiva y la justificación. Según el análisis aquí proporcionado, la memoria perceptiva se considera un tipo especial de creencia. Nuestra primera cuestión es, por tanto, bajo qué condiciones una creencia semejante está justificada. A ella contestamos con la respuesta, ahora familiar, del coherentismo. [cf. Brandt (1955).] Cualquier creencia está justificada por los efectos de su retención o su rechazo en el sistema global.

La segunda cuestión es la de justificar nuestra tendencia creciente a retener ciertas creencias por ser creencias de determinado tipo. ¿Por qué es una razón para conservar una creencia el que sea de ese tipo?

No podemos limitarnos a reiterar la respuesta que dimos en el caso de la percepción. Argüíamos allí que, en la percepción, nuestro contacto con el mundo es tan directo que, si tiene éxito, debe contar como conocimiento. Pero las rutas causales involucradas en la memoria perceptiva no son tan directas. No podemos decir que el suceso original, distante, puede ser la causa principal del recuerdo perceptivo actual. Muchas más cosas parecen necesarias, especialmente a medida que aumenta la distancia en el tiempo. Necesitamos encontrar algún otro medio de suponer que la confianza en la memoria perceptiva como tal está justificada.

Consideramos que la memoria fáctica es posible y que la memoria fáctica, en casos apropiados, es conocimiento. Quizá podamos utilizar esto para responder a la cuestión que nos ocupa. La memoria perceptiva es creencia presente causada por percepción pasada. Ya hemos aceptado que, en casos favorables, la percepción es conocimiento. De modo que, en casos favorables, la memoria perceptiva es creencia presente causada por conocimiento perceptivo pasado. Por tanto, alguien que se considere a sí mismo como teniendo memoria perceptiva debe considerarse a sí mismo como teniendo una creencia fundamentada en

conocimiento previo. En cuyo caso estará justificado al adscribir a esa creencia una dosis adicional de consistencia.

LECTURAS COMPLEMENTARIAS

D. Locke (1971) proporciona una revisión breve pero excelente de las teorías filosóficas de la memoria.

Russell (1921, cap. 9) muestra la desafortunada batalla de Russell con el problema del realismo indirecto.

Brandt (1955) rechaza las teorías coherentistas de la justificación, pero todavía se siente capaz de ofrecer un análisis de la justificación de las creencias-de-memoria que tiene un marcado aroma coherentista.

Dummett (1978, cap. 21) es un intento de dar el mejor sentido posible a la disputa entre realistas y antirrealistas sobre el pasado.

Martin y Deutscher (1966) intentan un análisis general de la memoria, subrayando la necesidad de una conexión causal entre la percepción pasada y la «representación» presente.

13. INDUCCIÓN

13.1. INDUCCIÓN, PERCEPCIÓN Y MEMORIA

En los capítulos anteriores hemos considerado nuestra capacidad de reunir conocimiento sobre nuestro medio circundante, y nuestra capacidad para retener y conservar ese conocimiento con posterioridad. En el capítulo presente consideraremos nuestra capacidad (o nuestra falta de capacidad) para ir más allá de tal conocimiento, para construir conocimiento nuevo sobre la base del conocimiento más antiguo. Podemos extender nuestro conocimiento por medio del razonamiento, al constatar que ciertas cosas que sabemos proporcionan razones a favor de otras creencias. Cuando esas razones sean suficientemente poderosas, es de esperar que, al creer de acuerdo con ellas, hayamos adquirido un conocimiento nuevo. Una creencia verdadera, basada en un conocimiento previo que proporcione justificación inferencial suficiente para esa creencia, es de esperar que sea conocimiento.

Hay dos formas de razonamiento, el deductivo y el inductivo. El razonamiento deductivo se da cuando consideramos que nuestra justificación inferencial es concluyente, en el sentido especial de que es imposible, so pena de contradecirnos, que las creencias que constituyen nuestras razones sean verdaderas y la conclusión que extraemos de ellas sea falsa. Cuando es cierto que sucede tal cosa, nuestro razonamiento deductivo es válido; en caso contrario, es inválido.

El razonamiento inductivo se da cuando consideramos que nuestras razones son suficientes para justificar una conclusión, sin ser concluyentes en el sentido anterior, o cuando pensamos que tenemos algunas razones, que no llegan a ser concluyentes, en favor de la conclusión, quizá a la espera de encontrar razones ulteriores, de modo que la suma total de razones sea suficiente. Esto puede expresarse más claramente en términos de probabilidad. Un argumento inductivo con éxito es el que convierte su conclusión en algo probable, o en algo más probable que cualquier alternativa igualmente detallada; la probabilidad (relativa) que da a su conclusión puede no ser suficiente todavía para justificar nuestra creencia en ella, porque puede haber razones

más poderosas en su contra, o porque el grado de probabilidad que se obtiene no es lo suficientemente grande para justificar nada más que una creciente tendencia a seguir buscando. Pero suponemos que es posible que, con razones adicionales, acabaremos por estar justificados para aceptar la conclusión.

El razonamiento inductivo se emplea en casi todas las ramas de la investigación humana. Podemos preguntarnos si lo que sabemos ahora sobre el presente proporciona justificación inductiva para las creencias sobre el pasado, como haría un detective en el curso de una investigación criminal. O podemos preguntarnos si nuestras creencias sobre el pasado justifican ciertas creencias sobre el futuro, como podría hacer alguien que estuviera considerando una oferta de matrimonio (aunque, en este caso, no creo que sea recomendable adoptar esta perspectiva). La inducción, por tanto, no está especialmente implicada en el conocimiento del futuro. Pero podría tentarnos la propuesta inversa: ¿es posible llegar al conocimiento del futuro de un modo que no sea la inducción?

Hay dos maneras en las que podríamos esperar llegar al conocimiento de los sucesos futuros que no son razonamiento inductivo. La primera es suponer que, a veces, sabemos lo que sucederá porque tenemos un conocimiento no inferencial de nuestras intenciones. Esto plantea cuestiones complejas que no podemos detenernos a considerar en detalle. Advirtamos sólo que no es obvio que esta forma de conocimiento del futuro sea completamente independiente de cualquier apoyo inductivo. Es posible que tengamos un conocimiento no-inferencial de nuestras intenciones, pero el conocimiento resultante sobre el futuro ¿no depende del conocimiento general de que, en ciertos asuntos, nuestras intenciones se llevan normalmente a cabo?

La segunda manera en la que podríamos esperar un conocimiento no-inductivo del futuro es la de suponer que, en ciertos casos, podemos ver lo que va a suceder. Esta sugerencia puede hacerse en casos extremos, como en los adivinadores de la suerte que observan sucesos futuros en una bola de cristal, o en casos más ordinarios en los que hablamos de ver que se producirá un choque o que la pelota saldrá fuera. Pero hay un argumento poderoso en favor de que en ninguno de estos casos podemos observar el futuro.

Es preciso que se trate de un argumento poderoso, dado que muchas de las cosas que se han dicho en los capítulos anteriores parecen dejar espacio para la idea de que, al menos en una forma limitada, la observación del futuro es algo posible. La creencia perceptiva puede ser asunto de grado. Algunas creencias son más perceptivas que otras,

algunas son muy perceptivas, algunas menos, y es posible que algunas no lo sean en absoluto. Esto nos deja espacio para decir, por ejemplo, que podemos ver que una escalada es difícil o que un acantilado es peligroso. No estamos obligados a decir que tales cosas sólo se pueden conocer por un razonamiento inductivo a partir de lo que puede decirse con más propiedad que es objeto de percepción. Es posible que haya la laxitud suficiente como para que no estemos obligados a eliminar todas las alusiones a nuestra percepción de lo que ha de suceder, que, de hecho, son habituales. Decimos, por ejemplo, que vemos que la pelota saldrá fuera.

El argumento crucial que nos lleva a eliminar completamente la percepción del futuro es un argumento sobre la dirección de la causación. Sugerimos con anterioridad que las creencias perceptivas difieren de otras en que los hechos que son sus contenidos pueden ser la causa principal de la creencia en ciertos casos favorables. Pero, para que haya percepción del futuro, las cosas que constituyen el contenido de la percepción deben estar en el futuro. ¿Y es posible que algo que está completamente en el futuro sea parte de la causa de un suceso actual? ¿Puede el futuro causar el presente? Incluso si concedemos que no es necesario que una causa preceda a su efecto, pudiendo ser contemporánea, hay argumentos a favor de que el presente no puede ser efecto de un suceso que todavía no ha ocurrido.

El ejemplo típico es el de alguien que se levanta invariablemente cinco minutos antes de que suene el timbre de su despertador, independientemente de la hora a la que lo haya conectado. ¿Qué impide que digamos que la causa de que se despierte es el hecho de que su despertador habría comenzado a sonar cinco minutos más tarde? Decir tal cosa sería afirmar que el efecto (el hecho de que se despierte) puede haber pasado antes de que suceda la causa. Lo que se supone que es imposible. Imaginemos que un día, habiendo puesto el despertador a las ocho en punto, se despierta a las ocho menos cinco y que el despertador no suena a las ocho por algún fallo mecánico. En una ocasión semejante, deberíamos buscar alguna otra explicación de que se hubiera despertado, y es probable que escogiéramos algún suceso anterior, como, por ejemplo, el que hubiera preparado el despertador para sonar a las ocho. Y si es esa acción anterior la que causa, en este caso, que se despierte posteriormente, ¿no deberíamos decir lo mismo en los demás casos? El que suene el timbre, algo que sucede normalmente cinco minutos después de que la persona se despierta, puede considerarse como un efecto distinto de una causa común, y no como una causa posterior de un suceso precedente.

Este argumento puede generalizarse. No podemos contestar diciendo que podría haber alguien que no se levantara cuando su despertador se hubiera estropeado y fuera a fallar. La razón está en que la estructura del ejemplo parecería exigir que el efecto hubiera pasado antes de comenzar la causa, o, al menos, que hubiera una brecha entre el comienzo del efecto y el de la causa durante la cual alguien podría intervenir para evitar que se diera la causa. No es necesario que nadie intervenga, basta con que tal intervención sea una posibilidad. Esta posibilidad es incompatible con la idea de que el despertarse está causado por el posterior sonido del timbre. Es esto lo que constituye que un suceso sea efecto de otro. B sólo es un efecto de A si el que se dé A es necesario para que se dé B. Para verlo, consideremos una mañana particular en la que nuestro hombre se despierta. Al suponer que tal suceso es el efecto del sonido del timbre, deberemos suponer que, si el despertador no fuera a sonar, nuestro hombre no se hubiera despertado. Pero el mero hecho de que podamos intervenir muestra que este condicional es falso. Porque se ha despertado y podría no sonar, sabemos que se habría despertado tanto si el despertador hubiera estado a punto de sonar como si no. Saber esto es saber que el que se despierta está causado por algo distinto al sonido del timbre.

La conclusión de este argumento, pues, es que el orden causal nunca puede ser el inverso del orden temporal [véase Flew (1954)]. Lo que demuestra, por la manera en la que la percepción necesita de una relación causal entre hecho y creencia, que no podemos percibir el futuro. Si el argumento nos ha impresionado, deberemos insistir en que no es realmente posible ver que lloverá o que la pelota saldrá fuera. Más bien lo que vemos es el tiempo que hace ahora o la trayectoria, y razonamos inductivamente a la lluvia futura o al camino que seguirá la pelota.

13.2. DOS CONCEPCIONES DEL FUTURO

Como antes, realistas y antirrealistas ofrecen concepciones alternativas del futuro. En esta área, el antirrealista considera que sólo llegamos a comprender las oraciones en tiempo futuro en circunstancias que aprendemos a considerar como evidencia de la ve/dad de esas oraciones, y que es increíble la sugerencia realista de que, con posterioridad, llegamos a comprender cómo podrían ser verdad tales oraciones en ausencia de las evidencias relevantes. En nuestra comprensión originaria de tales oraciones no hay dos elementos distintos que pudieran

ser diferenciados de ese modo. La oración se evalúa como verdadera o falsa de acuerdo con la evidencia pasada o presente. Por tanto, si no hay evidencia ni en favor ni en contra de la oración, no es ni verdadera ni falsa. La mayoría de las oraciones sobre el futuro serán de este tipo, por lo que carecerán de valor de verdad.

Esta posición tiene su atractivo. El futuro no ha sucedido todavía y tenemos la impresión de que no hay nada que puedan describir las oraciones en futuro, ni nada sobre el futuro que pueda hacer que, ahora, tales oraciones sean verdaderas o falsas. El carácter abierto del futuro, al haber un enorme número de hechos sobre el futuro de los que no tenemos esperanza alguna de conocimiento actual, no es un defecto epistemológico. Es una necesidad metafísica; lo que distingue el futuro del pasado. El pasado ya está cerrado, y su naturaleza hace que los enunciados en tiempo pasado tengan determinada (aunque quizá no sea determinable) su verdad o falsedad. No podemos decir lo mismo sobre el futuro.

Si encontramos que este tipo de contraste entre pasado y futuro es convincente, estaremos tentados por el antirrealismo respecto al futuro y por el realismo respecto al pasado. Un antirrealismo más completo rechazaría el contraste, y diría lo mismo respecto a nuestra concepción del pasado y del futuro. Dejando a un lado el argumento sobre el pasado, debemos considerar un argumento en favor de cierta forma de antirrealismo sobre el futuro. Se trata de la afirmación de que un realista sobre el futuro no puede conceder que tengamos libre albedrío.

Hay argumentos contra la posibilidad de la libertad de la voluntad que son independientes de esta controversia. El determinista causal argumenta que todo suceso y toda acción están causados por otros sucesos, y, por tanto, que no es verdad, en ningún caso, que hubiéramos podido actuar de forma distinta a como lo hicimos. Dado que lo que determina que un suceso cause otro es (o incluye) el hecho de que, dado el primer suceso y las circunstancias concurrentes, el segundo no podría haber sido distinto a como fue. Una acción libre es una acción que podría haber sido, en las mismas circunstancias, distinta a como ftie. Por tanto, si todas las acciones están causadas, ninguna es libre.

Este argumento determinista se ve discutido por el compatibilista, que argumenta que una acción puede ser, a la vez, causada y libre. No obstante, el argumento que nos interesa ahora no es el del determinista, sino el del fatalista. El fatalista supone, de algún modo, que ya está fijado lo que sucederá mañana, pero no por motivos causales. Sus razones son las del realista, que considera que toda oración sobre el futuro ya tiene ahora un determinado valor de verdad. Un fatalista

razonaría del siguiente modo: o bien moriré de un cáncer causado por fumar, o no. Si moriré de ese modo, no tiene sentido que deje de fumar. Si no, no tiene sentido que deje de fumar. Así que no tiene sentido que deje de fumar. Este razonamiento comienza con la asunción realista de que todo enunciado sobre el futuro tiene un valor de verdad; o es verdadero o es falso.

De una forma menos caricaturesca, podemos presentar del siguiente modo el argumento del fatalista contra la libertad de la voluntad. Ya es verdad que iré al restaurante el viernes o que no iré. Si ya es verdad que iré, ¿cómo puede decirse que tengo la opción de no ir, o viceversa? Parece como si el asunto ya estuviera fijado y nada pudiera cambiarlo. Sea cual sea la verdad, es inevitable y cualquier alternativa es imposible.

Hay suficiente peso en este argumento, a pesar de que parezca un mero truco, como para haber persuadido a Aristóteles de la necesidad de abandonar el realismo sobre el futuro (edición de 1962, cap. 9). Este movimiento se beneficia de las ventajas tradicionales en contra del escéptico, puesto que el antirrealista sostiene que lo que hace que un enunciado sobre el futuro sea verdad o es sólo la mejor de las evidencias que pudiéramos tener sobre su verdad. Por tanto, no hay ninguna posibilidad de que tengamos la evidencia relevante y el enunciado sea falso. Se rechaza como inconcebible la queja del realista de que siempre es posible, dada determinada evidencia, que el enunciado sea falso. De modo que el antirrealismo sobre el futuro es muy atractivo. La dificultad estriba en cómo adoptarlo sin adoptar una forma general de antirrealismo. Después de todo, ya hemos argumentado en contra del antirrealismo en teoría de la percepción.

Sin embargo, hay una forma más severa de argumento escéptico en esta área, que ataca, en general, la noción misma de evidencia en la que confía el antirrealista. Se trata de la cuestión de Hume sobre la inducción en general. Al considerarla, nos desplazamos de la mera cuestión del conocimiento del futuro hacia una preocupación más general sobre la posibilidad de conocer lo inobservado.

13.3. HUME Y SUS CRÍTICOS

Las cuestiones de Hume sobre el razonamiento inductivo ya se han planteado en 1.2; las hemos dejado descansar por algún tiempo y ahora se reavivan. No las repetiremos ahora, sino que empezaremos a considerar inmediatamente posibles respuestas a Hume.

¿ES VICIOSA LA CIRCULARIDAD?

Hume se queja de que el único argumento verosímil para justificar nuestro uso de la inferencia inductiva involucra un círculo vicioso, dado que apela a la experiencia para justificar las apelaciones a la experiencia. Pero algunos filósofos han sentido la tentación de mantener que, aunque haya algo como un círculo en una justificación inductiva de la inducción, no se trata de un círculo vicioso [por ejemplo, Black (1954), cap. 11, y (1958)]. La sugerencia es que un argumento como:

> El razonamiento inductivo se ha demostrado fiable en el pasado.
> Por tanto, el razonamiento inductivo es (normalmente) fiable,

tiene como conclusión el enunciado de que el principio de inferencia que nos lleva desde la premisa a esa conclusión es fiable. Pero no es una forma de circularidad; hay una diferencia crucial entre un principio de inferencia y un enunciado. Habría circularidad si el argumento requiriera realmente como una premisa la proposición que se comporta también como conclusión. Pero no lo requiere. Hay sólo dos razones para pensar que sí lo requiere, pero las dos son erróneas. Podríamos decir, en primer lugar, que el argumento sería inconcluyente de cualquier otro modo. Para convertirlo en concluyente, añadamos el enunciado relevante a las premisas; pero, en ese caso, el argumento será palmariamente circular. La respuesta a esto es que el argumento es un ejemplo perfectamente adecuado de razonamiento inductivo tal y como está. No hay ninguna necesidad de que añadamos premisas adicionales para hacerlo deductivamente válido, y, por tanto, concluyente. En general, no puede ser queja alguna contra los argumentos inductivos el que no sean deductivos; la justificación de la inducción no es el intento de mostrar que todos los argumentos inductivos son implícitamente deductivos.

En segundo lugar, podríamos pensar que, normalmente, un enunciado del principio de inferencia sobre el que descansa el argumento debe ser insertado, para que sea completamente explícito, como una premisa. Pero esta sugerencia lleva a un regreso al infinito, dado que el argumento resultante dependerá de algún principio de inferencia, que, por tanto, deberá insertarse; y el argumento resultante dependerá de un principio ulterior, etc.

Por tanto, se sostiene que el argumento no es implícitamente circular, porque no hay ningún sentido en el que la conclusión se necesite

como una de las premisas. Así que un argumento puede establecer la fiabilidad de su propio principio de inferencia cuando, como anteriormente, su conclusión asevera esa fiabilidad. Podemos justificar inductivamente la inducción.

Esto es ingenioso, pero no funciona. Para ver por qué, tratemos de considerar el asunto desde la perspectiva de Hume. ¿Nos da el argumento anterior alguna razón para aceptar su propia conclusión? Sólo puedo considerar que lo hace si acepto previamente sobre razones independientes, el principio de inferencia sobre el que descansa. De modo que el argumento nunca podría proporcionarnos una razón para aceptar su conclusión si no dispusiéramos previamente de razones suficientes para hacerlo.

APELACIONES A LA ANALITICIDAD

La cuestión de Hume es la de cómo podemos tener razón alguna para suponer que la observación pasada y presente proporciona evidencia de la que podamos inferir inductivamente. En ausencia de tal razón, insiste, no puede haber tal cosa como evidencia, tal cosa como la posesión de razones inductivas para creer.

Una respuesta clásica es sostener que no es posible cuestionar si la observación presente o pasada nos da evidencia o razones para creencias adicionales [véanse Edwards (1949) y Strawson (1952), cap. 9.2]. El enunciado de que la observación constituye evidencia es verdad a causa de lo que queremos decir por «evidencia». Alguien que dudara si la órbita observada de un planeta era evidencia sobre su órbita futura mostraría, con ello, que no comprendía el significado de la palabra «evidencia». No hay ninguna posibilidad de que pudiéramos equivocarnos sobre el significado de «evidencia», y no hay, por tanto, ninguna posibilidad de que esas observaciones dejaran de proporcionar evidencia sobre lo inobservado.

Esto se denomina la justificación analítica de la inducción, porque equivale a sostener que el enunciado «el pasado observado es evidencia para el futuro» es analítico; es verdad en virtud de los meros significados de sus palabras. Como tal, por supuesto, no es accesible directamente a aquellos que, con Quine, rechazan la noción de analiticidad y el contraste tradicional entre analítico y sintético. Pero hay otro argumento adicional contra la justificación analítica, que puede ser atractivo a quineanos y no quineanos. Se debe a Urmson (1953).

Este otro argumento parte de la observación de que decir de algo

que es evidencia es, en parte, evaluarlo, considerarlo como una guía fiable que estamos justificados para seguir. Ahora bien, los términos que se usan para esta evaluación tienen una característica peculiar (aunque es posible que esta característica sea menos notoria en otros casos) que puede verse más claramente en el caso del término de aprobación más general «bueno». Aprendemos el uso de este término por apelación a ejemplos como los que nos presentan nuestros mentores (padres, tutores escolares...). Pero ello no nos obliga a limitarnos a aprobar sólo, o todas, las cosas que ellos aprueban. Sea como sea, nuestra comprensión del término «bueno» puede dejar de depender de los ejemplos originales. Podemos llegar a aprobar conjuntos de objetos radical o completamente diferentes, o ninguno en absoluto, sin que por ello mostremos que hemos olvidado lo que se nos enseñó. Y lo mismo es verdad con la palabra «evidencia». Nuestra comprensión de esa palabra, indudablemente obtenida en circunstancias normales, es tal que, sin abusar de ella, podemos llegar a considerar tipos diferentes de cosas como evidencia, o incluso preguntarnos si existe algo que sea realmente evidencia para algo más. Esto es, por supuesto, lo que hizo Hume. Parece, pues, que la justificación analítica no tiene éxito al tratar de dejar a un lado, como incoherente, la cuestión de Hume.

Podría objetarse que no hemos advertido toda la fuerza de la justificación analítica. El antirrealista podría pretender que la justificación analítica es una respuesta antirrealista a la cuestión realista; y, al habernos limitado a afrontarla desde una perspectiva realista, habríamos dejado a un lado lo esencial. Es posible que Hume suponga que hay asuntos de hecho, sobre el pasado inobservado y sobre el futuro, sobre los que la experiencia acumulada se considera, normalmente, evidencia relevante aunque no concluyente. A continuación señala que la experiencia no puede proporcionarnos razón alguna para suponer que podemos cruzar la brecha entre lo observado y lo inobservado. La réplica antirrealista es insistir en que no hay tal brecha entre los asuntos de hecho y las proposiciones que consideramos que constituyen evidencia relevante. Nuestra comprensión de qué es que una proposición sobre lo inobservado sea verdadera está indisolublemente vinculada a los tipos de consideración que aceptamos como evidencia concluyente. La razón por la que es una verdad analítica que el pasado observado es evidencia para el futuro es, pues, la afirmación general de que nuestros conceptos de verdad y evidencia van a la par. El que una proposición sobre el futuro sea verdadera es, exactamente, que exista evidencia accesible de que es verdadera. Comprender qué es que un enunciado sea verdadero es, exactamente, saber lo que cuenta

como evidencia de su verdad. Por tanto, no podemos suponer que estamos equivocados al considerar que ciertas observaciones son evidencia relevante. Que sean evidencia relevante está determinado no tanto por el significado de la palabra «evidencia», como por el significado de las proposiciones respecto a las que cuentan como evidencia.

La versión antirrealista de la justificación analítica de la inducción parece más impresionante. Para poder elaborar una respuesta adecuada a ella es preciso que nos aproximemos desde un punto de vista un tanto diferente.

13.4. GOODMAN: EL NUEVO ROMPECABEZAS DE LA INDUCCIÓN

Hume planteó una cuestión en términos de la práctica de inferir, a partir de regularidades observadas, la probable continuidad de esas regularidades. Argumentó que la práctica misma de inferir no podría justificarse, si tal justificación requiriera razones para creer que el razonamiento inductivo es fiable. Pero no concluyó de ello que la *práctica* fuera irracional. Sugirió que la naturaleza humana es tal que adquirimos ciertas expectativas habituales después de observar una regularidad suficiente en la naturaleza. No tenemos ninguna razón para razonar inductivamente, pero no podemos dejar de hacerlo. Dado que comprendemos qué es ser racional no por inferencia a lo que debemos hacer (y normalmente no hacemos), sino por referencia a lo que de hecho hacemos, la inferencia inductiva es una práctica racional, aunque no esté fundamentada en razones.

Nelson Goodman argumenta que la solución de Hume plantea, de hecho, cuestiones de tipo similar que son incluso más difíciles (1973). La respuesta de Hume es que la observación de patrones regulares produce en nosotros ciertas expectativas habituales. Observamos tan a menudo que los objetos caen cuando se les suelta, que esperamos naturalmente que el próximo haga lo mismo. Y la inferencia inductiva correcta se define en términos de inferencia a sucesos similares a los observados. La inferencia más fiable se supone que es aquella cuya conclusión sugiere que el mundo continuará siguiendo el curso más semejante al que ha seguido hasta ahora. Pero Goodman afirma convincentemente que esta apelación a semejanzas oculta un supuesto que tiene difícil justificación.

Supongamos que, hasta este momento, todas las esmeraldas observadas han sido verdes. Los cánones de la inferencia inductiva, tal y

como se explicitaron anteriormente, nos permiten inferir que la próxima esmeralda será verde. Pero esta historia dejará de ser convincente si suponemos que hay otro predicado «verdul», que tiene el siguiente sentido: un objeto es «verdul» en el momento del tiempo *t* si y sólo si o bien es verde y *t* es anterior al 1 de enero del año 2000, o bien es azul y *t* es posterior al 1 de enero del año 2000. Dado este predicado, cualquier evidencia a favor de que las esmeraldas futuras han de ser verdes estará también a favor de que hayan de ser azules. La razón es que las esmeraldas observadas no eran más verdes que «verdules». Es tan verdad que son verdes como que son «verdules», por lo que estamos tan justificados para concluir que las esmeraldas serán azules en el futuro como para concluir que serán verdes.

Lo que quiere decir que la inferencia inductiva correcta no puede caracterizarse en términos de inferencia de la continuación de semejanzas previamente observadas. Estamos inclinados a suponer que una de las inferencias sobre las esmeraldas es correcta, y que la otra no lo es (y, por supuesto, hay un número infinito de inferencias semejantes, definiendo predicados como «vermarillo», «verrojo», y aún peores). Pero todavía no hemos dado razón alguna para preferir una a la otra. Ni tenemos un análisis aceptable de qué sea una inferencia inductiva correcta, por apelación al cual pudiéramos confiar en que el uso de tales inferencias estuviera justificado.

La respuesta natural es protestar diciendo que hay algo profundamente sospechoso con relación a predicados artificiales como «verdul». Pero no hay consenso respecto a qué pueda haber de equivocado en ellos. (En cierto modo, el nuevo rompecabezas es simplemente la cuestión de por qué son incorrectos.) El mero hecho de que sean artificiales no fundamenta ninguna queja sobre ellos; sólo muestra que no lo usamos, no que no los deberíamos usar.

Una respuesta común es la de decir que un predicado «sensato» no puede contener ninguna referencia a un punto del tiempo (o del espacio) particular. «Verdul» se define en términos de verde antes de *t* y, por tanto, no es sensato. Pero hay dos errores en este argumento. En primer lugar, es igualmente cierto que «verde» se define en términos de su «verdulez» antes de *t*; un objeto es verde en *t* si y sólo si es «verdul» en *t* y *t* es anterior al 1 de enero del año 2000, o es «azude» en *t* y *t* es posterior a esa fecha. «Azude» se define como «verdul», pero intercambiando en la definición los términos «azul» y «verde». De modo que, de acuerdo con este criterio, todos los predicados parecen igualmente insensatos, desde el punto de vista de quien utilice otros. En segundo lugar, incluso si debiéramos denominar «insensatos» a

tales predicados, todavía no se nos ha dicho qué es lo que hay en ellos que hace que las inferencias que los utilizan no sean fiables.

En vez de proseguir con otros intentos de solucionar el rompecabezas de Goodman, terminaremos relacionándolo con otras de nuestras preocupaciones anteriores. En primer lugar, la respuesta antirrealista a Hume que se ha ofrecido anteriormente no nos proporciona nada que podamos aprovechar en contra del rompecabezas de Goodman. El antirrealista apela al «hecho» de que usamos las observaciones pasadas como evidencia para los casos futuros, o que nuestra comprensión de qué sea la verdad de algo inobservado está vinculada a lo que consideramos como mejor evidencia a favor de esa verdad. Nuestra comprensión del hecho de que algo sea verde en el futuro está vinculada a lo que consideramos como evidencia de que será verde; así pues, dado el significado del enunciado relevante, no hay ninguna posibilidad de que estemos equivocados en lo que consideremos evidencia para él. Pero estas observaciones no discriman entre «verde» y «verdul». La única esperanza de que lo hicieran así reside en la constatación de que hemos estado usando el predicado «verde» y no el predicado «verdul». Y ésta es una idea de dudosa utilidad, por dos razones. En primer lugar, no está claro qué importa el predicado que hayamos usado; según este análisis, no vamos a poder mostrar que el otro es incorrecto. En segundo lugar, y más importante, ¿qué es lo que, en nuestra práctica anterior, determina que de hecho hayamos usado «verde» y no «verdul»?

Es verdad que hemos usado la *palabra* «verde». Pero ¿qué demuestra que no hemos estado pensando en términos de «verdulez» todo el tiempo? ¿Qué demuestra que no vamos a denominar «verdes» a los objetos azules después del 2000? Hasta que no digamos qué concepto hemos estado usando, no podemos tener esperanza alguna de argumentar en favor de uno de ellos sobre la base de que es el que hemos usado.

Esto es importante porque una razón atractiva para dejar de lado conceptos como «verdulez» es que no pueden ser adquiridos a través de ejemplos [Small (1961)]. La «verdez» es un concepto que podemos aprender por medio de ejemplos, como sabemos por haberlo adquirido de ese modo. Pero este argumento se viene abajo cuando nos enfrentamos a la posibilidad de que el concepto que hayamos estado usando todo el tiempo sea el de «verdulez». Si tenemos tal concepto, lo debemos haber adquirido a través de la experiencia de objetos «verdules».

La cuestión a la que hemos llegado, y el intento de responderla por medio de ejemplos, debe recordarnos en gran medida las observaciones de Wittgenstein sobre el seguimiento de reglas (5.5.-7). La cuestión de Goodman es, y debe ser, semejante a la de Wittgenstein, dado

que Goodman se pregunta qué es lo que justifica una manera de actuar más bien que otra. Y ésa es, precisamente, la cuestión que se planteaba Wittgenstein.

13.5. EL COHERENTISMO Y LA INDUCCIÓN

La situación actual es que tenemos dos formas de escepticismo inductivo con las que vérnoslas, la de Hume y la de Goodman. Como vimos, el antirrealismo proporcionaba algún tipo de respuesta a Hume, pero fracasaba por completo a la hora de enfrentarse al rompecabezas de Goodman. Veremos ahora qué puede aportar el coherentismo.

Los coherentistas hacen afirmaciones ambiciosas sobre la capacidad del coherentismo para proporcionar una perspectiva dentro de la cual se venga abajo el escepticismo sobre la inducción. Ewing sostuvo que «el principio de coherencia proporciona la única justificación racional para la inducción» [Ewing (1934), p. 247]; Blanshard está de acuerdo [Blanshard (1939), vol. 2, pp. 504-505]. Hay dos restricciones dentro de las cuales evaluar esas afirmaciones. La primera es nuestra preferencia presente por el internalismo; las respuestas del externalismo son insuficientes. La segunda es que la respuesta debe ser tan efectiva contra Goodman como contra Hume.

Supongamos que tenemos un breve enunciado de un principio inductivo de inferencia, PII. Un movimiento externalista sería decir que la adopción o el uso de PII tiene como resultado un incremento de la coherencia en el propio conjunto de creencias; esto es lo que justifica nuestro uso de PII. Un internalista añadiría que creemos que la adición de PII incrementaría la coherencia, y que, en la medida en que esa creencia sea verdadera, estamos justificados para usar PII. El internalismo no tiene que mostrar que su creencia verdadera está justificada, tal y como concedimos al considerar los grados de internalismo en 9.3; tal cosa nos conduciría a un regreso de justificaciones, y, por ello, nos arrojaría directamente en brazos de Hume.

Pero todavía está la cuestión de si es verdad que la adopción de PII incrementa siempre la coherencia. No tiene ningún sentido que intentemos mostrar tal cosa apelando al pasado; lo que involucraría una forma dañina de circularidad. Por tanto, la cuestión de si el uso de PII lleva a un incremento de la coherencia no puede ser completamente empírica. Necesitamos una razón por anticipado, un vínculo conceptual entre la inferencia inductiva y el incremento de la coherencia explicativa.

El coherentismo puede proporcionar ese tipo de razón. Lo hace así, manteniendo que el principio conductor de la inferencia inductiva es, exactamente, la «inferencia a la mejor de las explicaciones alternativas» [Harman (1970), p. 89)]. Por ejemplo, un detective infiriendo la culpabilidad de un sospechoso a partir de ciertas evidencias trata de llegar a la hipótesis que proporcione la mejor explicación de toda la evidencia. No es, por ello, un accidente que la inducción lleve a un incremento en la coherencia explicativa. Y, dado que tanto la verdad como la justificación se ven en términos de coherencia, podemos decir que el uso de la inducción debe aproximarnos a la verdad. ¿Qué otra mejor justificación de la inducción podríamos esperar?

Y hay más que esto. Todavía no hemos explicado la afirmación de Ewing de que el coherentismo es la *única* posición con éxito en esta área. Para considerar las bases de esta afirmación, necesitamos rehacer los pasos de Hume.

¿Cómo podemos saber o estar justificados en cualquier creencia sobre lo que sucederá a continuación? Vemos que un ladrillo va a romper la ventana; nuestra creencia natural sobre el futuro se alcanza por medio de la inferencia inductiva. ¿Qué justifica tal inferencia? No hay ninguna conexión necesaria que vincule los sucesos diferentes del vuelo del ladrillo y del rompimiento del cristal. No hay ninguna contradicción en suponer que se da el uno y no el otro, como sucedería en el caso de que estuvieran conectados por necesidad lógica. Y no hay ninguna otra noción comprensible de necesidad que la de necesidad lógica. Siendo así, esos dos sucesos van unidos sin estar conectados. Nuestra inferencia de uno al otro debe derivarse, por tanto, de la experiencia de conjunciones similares en el pasado. Y este tipo de inferencia no puede estar justificado empíricamente sin circularidad (ver 1.2).

Los coherentistas como Ewing y Blanshard detienen esta cadena al negar uno de sus eslabones cruciales. Sostienen que hay otra noción comprensible de necesidad, necesidad natural, que vincula conjuntamente los sucesos individuales, y mantienen que cualquiera que niegue tal cosa está condenado al escepticismo sobre la inducción, por lo que su punto de vista ofrece la única esperanza.

El tipo de necesidad natural del que hablamos es básico para la posibilidad de la explicación; sin ella la explicación sería imposible. Explicar algo es ver por qué, en esas circunstancias, debiera suceder. Los filósofos en la tradición de Hume, Hempel, por ejemplo (1942), consideran que la explicación se da cuando estamos convencidos de que el suceso relevante va o iba a suceder. Pero esto no basta. A menos que podamos ver *por qué* va a suceder, no tenemos ninguna explica-

ción. Y, para ver por qué, necesitamos algo más que el conocimiento de que en casos previos y semejantes sucedió tal tipo de cosa. Este tipo de conocimiento puede capacitarnos para hacer predicciones arriesgadas, pero no nos proporciona ninguna explicación; para la explicación necesitamos comprensión (cf. 11.1). El argumento de Hume es que, dado que no hay nada como una relación de necesidad entre sucesos, estamos limitados al tipo de explicación que él nos ofrece. Pero la respuesta es que, si la explicación es posible en absoluto, debe existir una relación necesaria entre los sucesos. En el caso simple del ladrillo y la ventana, no se trata sólo de que la ventana se va a romper; debe romperse, está obligada a ello. Nuestro conocimiento inductivo de que se romperá es conocimiento inferencial. El vuelo del ladrillo *entraña* que la ventana se romperá; puesto que, dado el ladrillo, la ventana tiene que romperse. Y sabemos tal cosa porque sería mucho más difícil explicar que la ventana no se rompiera que explicar su rompimiento. (Hay, de hecho, vínculos de explicación mutua.)

Ewing y Blanshard pretenden escapar al argumento de Hume negando su atomismo, el punto de vista de que los sucesos individuales están conjuntados, pero no conectados. Es el atomismo el que origina el escepticismo sobre la inducción, más bien que la naturaleza del caso.

Pero esto todavía nos deja con Goodman. Goodman señala que, tanto como nuestra práctica inductiva, hay muchísimas otras, cada una consagrada con su propio lenguaje; y no parece que haya nada para escoger entre los lenguajes. Cada práctica goodmaniana es igualmente inductiva. De modo que, si una está justificada, lo están todas. Supongamos por el momento que el argumento anterior sobre Hume nos permite aceptar la conclusión de que nuestra práctica inductiva está justificada. ¿Por qué debe preocuparnos el hecho de que otras prácticas estarían igualmente justificadas? Después de todo, el hecho de que muchos conjuntos de creencias estarán tan justificados como el nuestro no nos preocupa, la justificación puede ser compartida. La objeción de la pluralidad (8.2) sólo hace daño cuando consideramos pretensiones de verdad, más bien que de justificación. De modo que ¿no es igualmente aceptable el tipo de pluralidad que señala Goodman?

Lo que necesita mostrar Goodman es que, dado que otras prácticas son posibles, la nuestra no está justificada. Si la nuestra lo está, las otras lo están también. Y el punto crucial es que, por cada creencia que estemos dispuestos a sostener, habrá otra práctica, tan bien fundamentada como la nuestra, que recomiende la creencia opuesta. Por

ejemplo, el uso del predicado «verdul» nos lleva a concluir que las esmeraldas no son verdes sino azules después del 1 de enero del año 2000. Lo que quiere decir que las dos prácticas no son sólo diferentes, sino alternativas opuestas. Se oponen y nada justifica la adopción de una más bien que la de la otra, de modo que nada justifica la creencia de que las esmeraldas continuarán siendo verdes más bien que la creencia de que no lo serán.

En este momento, debemos buscar ayuda en la analogía de Wittgenstein. Lo que hemos descubierto es que no hay ningún rasgo interno de nuestra práctica que no pueda convertirla en algo más justificado que sus alternativas. Tampoco hay ningún rasgo externo que pueda hacerlo. El único factor relevante era que nuestra práctica es la que usamos, mientras que las otras no. Esto no ayuda demasiado, pero no debemos desesperar. La conclusión de los pensamientos de Wittgenstein sobre seguimientos de reglas era que una práctica no necesita de un fundamento externo independiente a modo de justificación (5.6). La razón por la cual buscamos ese fundamento externo es que pensamos que tales cuestiones son cuestiones externas, que nos preguntan sobre nuestra práctica desde un punto de vista que se supone externo a ella. Pero no lo son. Del mismo modo que las cuestiones sobre la justificación de alguna parte de nuestro uso del lenguaje, por ejemplo, nuestro uso de los predicados morales, son cuestiones que se preguntan en el seno de nuestra práctica lingüística. De modo que las cuestiones sobre la inducción, que constituye una forma de vida omnipresente, que parece casi constituir nuestro paradigma de conducta racional, son todavía cuestiones que se preguntan y deciden en el seno de nuestra visión del mundo. Hay que admitir que no pueden justificarse desde el exterior. Pero tal cosa no debe desconcertarnos. Si insistimos en que las cuestiones sobre justificación sólo pueden plantearse desde el seno de nuestra práctica, sentiremos menos vergüenza de las respuestas que tanto nos tientan.

Si esta respuesta wittgensteiniana a Goodman es efectiva, nos damos cuenta de que el movimiento de Goodman, al ir más allá de Hume, no era más que el desplazamiento desde una cuestión que tiene respuesta, en términos coherentistas, a una cuestión que no tiene posibilidad de respuesta. Sería una cuestión que nadie podría responder en el sentido en que se plantea, que conlleva su propia imposibilidad de encontrar respuesta y que, por tanto, se sitúa fuera del juego. Si podemos dejar a Goodman a un lado, Ewing tenía razón al afirmar que el coherentismo es la única posición desde la que puede mostrarse que la inducción es una práctica racional.

LECTURAS COMPLEMENTARIAS

Hume (edición de 1955, cap. 4.2) y Goodman (1973, caps. 3-4) ofrecen dos argumentos clásicos.

Las respuestas a Goodman incluyen las de Barker y Achinstein (1960), Small (1961) y Quine (1965, cap. 5).

Swinburne (1974) es un buen conjunto de respuestas a Hume, incluyendo las de Russell, Edwards y Black. La introducción es una visión útil y general.

El argumento contra la causación retroactiva está en Flew (1954). Dummett (1978, cap. 18) muestra grandes dosis de originalidad.

Cahn (1967) constituye una valiosa discusión de varios intentos (el de Aristóteles entre ellos) de enfrentarse al fatalismo.

Una buena formulación del determinismo se halla en Van Inwagen (1975). El compatibilismo de Hume se encuentra en Hume (edición de 1955, caps. 7-8).

Hay discusión sobre si el problema humeano de la inducción debe verse como abiertamente escéptico, más que como un intento de reducir al absurdo una teoría incorrecta de la racionalidad. Véase Stroud (1977, cap. 1).

Harré y Madden (1975, caps. 3-4) discuten las conclusiones de Hume. La respuesta coherentista a Hume la dan Ewing (1934, cap. 4.3) y Blanshard (1939, vol. 2, pp. 504-511).

14. CONOCIMIENTO *A PRIORI*

14. 1. EL FUNDAMENTALISMO Y EL CONOCIMIENTO *A PRIORI*

Hasta ahora nos hemos ocupado, casi exclusivamente, del conocimiento empírico, del conocimiento que sólo puede adquirirse por medio de la experiencia. En este capítulo consideraremos la naturaleza, el alcance y el objeto de otra supuesta forma de conocimiento, el conocimiento *a priori*.

El fundamentalismo distingue entre la justificación inferencial y la no inferencial. Nuestras creencias no básicas están justificadas inferencialmente por apelación a las básicas, que se justifican de un modo diferente y no inferencial. Pero ¿qué es lo que justifica los principios de inferencia sobre los que debe descansar cualquier justificación inferencial? Tales principios no son justificables en sí mismos por apelación a las creencias básicas, ni como conclusiones de inferencias a partir de esas creencias. Por tanto, para que el fundamentalismo tenga éxito, deberá existir una tercera forma de justificación de los principios de inferencia.

Como hemos visto (4.1), el fundamentalismo es una forma de empirismo, y su distinción entre creencias básicas y no básicas trata de reflejar la distinción entre aquellas creencias que vienen dadas por la experiencia y aquellas otras que derivamos por medio de inferencias a partir de la experiencia. [Esta distinción se hace en Russell (1926), p. 226]. El fundamentalismo, por tanto, nos proporciona una razón empirista para aseverar la existencia de algún conocimiento que no sea empírico. Pero también hay otras razones.

14.2. EL EMPIRISMO, LO *A PRIORI* Y LO ANALÍTICO

El empirismo tradicional, en su forma extrema, mantenía una distinción entre ideas y proposiciones. Las ideas son partes componentes de las proposiciones; las proposiciones son complejos de ideas, combinadas de tal manera que pueden ser verdaderas o falsas. (Las ideas

en sí mismas no se pueden caracterizar como verdaderas o falsas.) Los empiristas radicales sostuvieron que todas nuestras ideas se derivan de la experiencia, y que no puede saberse que ninguna proposición o combinación de ideas es verdadera sin apelación a la experiencia.

La primera de estas pretensiones no se discutirá aquí, aunque sólo sea porque hacerlo así sería aceptar la distinción entre ideas y proposiciones; encontraremos muy pronto razones para sentirnos incómodos con esa discusión. John Locke intentó establecerla de la única manera que pudo concebir, examinando todas las ideas cruciales, una a una, y mostrando cómo podrían derivarse de la experiencia. En una batalla denodada, alcanzó éxitos y fracasos. Los casos difíciles incluyen las ideas de identidad, igualdad, perfección, Dios, poder y causa.

La segunda de estas pretensiones se abandonó tan pronto como fue formulada, porque entraña una contradicción con el supuesto, altamente verosímil, de que podemos conocer que muchas proposiciones son verdad sin ningún otro uso de la experiencia adicional al que se requiere para la misma adquisición de las ideas (o conceptos) relevantes. Este supuesto es verosímil por el número de ejemplos que le dan apoyo. «Rojo es un color»; «2 + 3 = 5»; «un cuñado es el cónyuge de un hermano», y (un ejemplo más complicado del propio Locke) «donde no hay propiedad no puede haber injusticia». No necesitamos de la experiencia para verificar esas verdades; cualquiera que las entienda (que tenga las ideas relevantes) ya está en situación de conocer su verdad sin necesidad de investigación empírica.

Sin embargo, nada de esto debe perturbar, ni perturbó de hecho, a los empiristas. Aceptaron que tenemos conocimientos que no se obtienen a través de la experiencia. Lo hicieron al aceptar que las proposiciones implicadas gozan de un *status* especial. Todas ellas expresan relaciones entre nuestras ideas y nuestros conceptos, siendo el conocimiento que de ellas tenemos un *conocimiento conceptual*, sobre los conceptos, más que conocimiento sustancial, no conceptual, sobre el mundo real. Locke estaba de acuerdo en que estas verdades conceptuales (que consideró verdades verbales) podían conocerse por el ejercicio de la razón más que del de la experiencia, pero sostuvo que tal cosa era compatible con, y explicable por, su empirismo sobre los conceptos. Las ideas cuyas interrelaciones podemos conocer de ese modo son todas ideas empíricas.

El empirista que acepte que hay conocimientos *a priori*, por tanto, debe proporcionar una explicación de él en sus propios términos. Locke sostuvo que las proposiciones que se pueden conocer *a priori* son proposiciones triviales que expresan relaciones entre nuestras

ideas. Los empiristas del siglo XX, como Ayer, han defendido que una proposición sólo puede ser conocida *a priori* si es *analítica*, es decir, verdadera en virtud de los significados de sus palabras, más que en virtud de la forma de ser del mundo. Según este punto de vista, todo el conocimiento *a priori* es de verdades analíticas; las verdades sintéticas sólo pueden conocerse empíricamente. Por ello, se pensó que la aceptación del conocimiento analítico no podía ser dañina para el programa empirista, dado que tal conocimiento no pretendía ser conocimiento sustancial sobre el mundo. El empirismo se convirtió en el punto de vista menos radical de que toda verdad sustancial sobre el mundo es empírica; sólo las verdades analíticas pueden ser conocidas *a priori*.

Este resultado no se ajusta muy bien a la tendencia fundamentalista a mantener que nuestro conocimiento de los principios de inferencia es *a priori*, ya que tales principios son escasamente analíticos. No son verdaderos en virtud sólo de los significados de las palabras, como vimos en el caso de la inducción (13.3). Parecen ser sintéticos. En ese caso, ¿cómo puede decir el empirista que los conocemos *a priori*? Pero ahora no estamos defendiendo la doctrina empirista; nuestra tarea, por el momento, es la explicación.

14.3. ¿ES POSIBLE CONOCER *A PRIORI* VERDADES SINTÉTICAS?

¿Hay verdades de las que no puede decirse de un modo convincente que sean triviales o analíticas y que, sin embargo, sean conocidas *a priori*? Ya hemos considerado un ejemplo posible de lo sintético *a priori*, o sea, nuestro conocimiento *a priori* de los principios sintéticos de inferencia. Hay otros ejemplos.

Uno sería el de la exclusión de colores. Sabemos que ningún objeto es completamente verde y completamente rojo en el mismo aspecto y al mismo tiempo. Y lo sabemos sin ninguna necesidad de comprobarlo por apelación a la experiencia, porque sabemos por anticipado que no es posible tener una experiencia como de algo que sea completamente rojo y completamente verde a la vez.

Podríamos objetar a esto señalando que, seguramente, es un hecho de experiencia que ser rojo excluye ser verde: en algún sentido, puede ser algo que podamos verificar sin ulterior experiencia, una vez que ya sepamos qué es ser verde y qué es ser rojo. Pero el hecho de que un

color excluya otro no es un hecho sustancial sobre el mundo, por lo que nuestro conocimiento de que sucede así no es el tipo de conocimiento de verdades sintéticas *a priori* que estamos buscando. En vez de ello, es un simple hecho empírico sobre la *apariencia* que nada parece tener simultáneamente dos cualidades «competidoras».

Esta réplica no funcionará, por dos razones. En primer lugar, es sentar la cuestión por anticipado pretender que podemos conocer por medio de la experiencia que dos cualidades compiten del modo relevante. La experiencia nos dice que ningún objeto es, a la vez, completamente rojo y completamente verde. Pero ¿qué hay en la experiencia que nos revele el hecho ulterior de que ningún objeto *puede* ser de esa forma? Sería más plausible decir que la exclusión de colores no se revela por la experiencia, sino que se origina simplemente por nuestra manera de distinguir entre diferentes colores. Distinguimos de tal manera entre rojo y verde que no permitimos que nada cuente como algo rojo y verde a la vez. De modo que la exclusión de los colores no es más que un artificio de nuestro esquema conceptual, no un reflejo de la manera en que el mundo se comporta (y se debe comportar).

Pero ¿qué hay que decir de la intuición de que no inventamos la exclusión de colores para nuestros propios propósitos, sino que nos limitamos a reflejar en nuestro uso del lenguaje (o de nuestro esquema conceptual) las relaciones de exclusión que descubrimos en el mundo? Es posible que pudiera argumentarse, en el caso de los colores, que el color es asunto de apariencia y, por tanto, una relación entre nosotros y el mundo, por lo que la exclusión de colores es tanto un hecho sobre nosotros —sobre el modo como nos parecen las cosas— como sobre el mundo que nos rodea. Esto es reformular la posición ya planteada en los parágrafos precedentes.

La segunda objeción a esta posición es que hay otros ejemplos del tipo de relaciones de exclusión que ahora nos ocupa y de los que no puede decirse que traten de la *apariencia* en el mismo sentido. Por ejemplo, ¿qué sucede con el hecho obvio de que una vaca no es un caballo? Nada puede ser, a la vez y respecto a lo mismo, caballo y vaca. Y este hecho, del tipo que sea, no es un hecho sobre las apariencias. Ni es muy convincente decir que es un hecho sobre nuestro esquema conceptual, más bien que sobre vacas y caballos.

¿Es posible que sea parte del significado de «es una vaca» que las vacas no son caballos, por lo que «vaca» implica, entre otras cosas, «no es un caballo»? Si así fuera, la verdad de que las vacas no son caballos sería una verdad analítica, y no sería un caso de verdad sintética conocida *a priori*. Pero, si esto fuera parte del significado de «es

una vaca», es de presumir que cualquiera que conociera ese significado sabría también que las vacas no son caballos. ¿Y no podría darse el caso de que alguien no supiera nada sobre caballos, y tampoco conociera el significado de «es un caballo», pero que conociera perfectamente bien el significado de «es una vaca»? Para persuadirnos de esto, sólo es preciso que recordemos los miles de otras cosas que no son las vacas: televisiones, mariposas y gazpacho, por ejemplo. Si el significado de «es una vaca» contuviera, de algún modo, todas esas relaciones de exclusión, sería imposible que nadie lo aprendiera, en particular, cuando recordemos que las otras palabras contendrían también sus propias relaciones de exclusión. De modo que la proposición de que nada es a la vez vaca y caballo no es analítica. Pero no es sobre apariencias como se podría sostener de las proposiciones correspondientes sobre colores. Y nuestro conocimiento de ella no puede, según ese análisis, considerarse empírico. Parece que, si la conocemos, la conocemos *a priori*, porque no necesitamos, para verificarla, de ninguna experiencia adicional a la que necesitamos para saber qué es ser una vaca y qué es ser un caballo. Por tanto, éste parece ser un caso de conocimiento *a priori* de una verdad sintética.

Un argumento similar puede introducirse con relación al problema, más difícil incluso, de nuestro conocimiento de la verdad matemática. Si concedemos que el conocimiento matemático es *a priori*, ¿qué vamos a decir de las verdades matemáticas? ¿Son analíticas o sintéticas? «7 + 5 = 12» ¿es verdadera en virtud del mero significado de sus símbolos? Kant utilizó la palabra «analítico» de un modo ligeramente diferente al que se ha venido utilizando aquí, definiéndola no en términos de significado, sino de conceptos. Sostuvo que un juicio es analíticamente verdadero si el concepto que aparece como predicado está contenido en el concepto que aparece como sujeto; un buen ejemplo sería «un hermano es de sexo masculino» (Kant, edición de 1961, Introducción 4; B10-14). En este sentido de «analítico», mantuvo por algún tiempo que las verdades simples de la matemática son analíticas, pero más tarde lo puso en duda (Kant, edición de 1967, pp. 128-131). Por ejemplo, 12 es la suma de 5 y 7, y podríamos sentir la tentación de decir que el conocimiento de tal cosa es un requisito de la posesión del concepto de 12. Pero 12 es también el resultado de infinitas sumas de otros números, el producto de otros, etc. El conocimiento de cualquiera de estos hechos tendría el mismo derecho a ser considerado como un requisito de que alguien poseyera el concepto de 12, y, como no podemos admitir a todos ellos, no podemos admitir ninguno. (La única de estas verdades de la que sería posi-

ble decir que tiene un *status* especial sería «12=11 + 1», que definiría 12 como el sucesor de 11.) Por tanto, la verdad matemática es sintética, en el sentido de Kant.

El mismo argumento sería también efectivo si se formulara con nuestra definición de «analítico» en términos de significado. Del mismo modo que hay demasiadas propiedades excluidas por el hecho de que «es una vaca» se satisfaga, hay demasiadas interrelaciones entre un número y otros números para que todas ellas estén incluidas en el significado de «12». Por tanto, la verdad matemática es sintética, en nuestro sentido, aunque sea cognoscible *a priori*.

Hasta ahora el argumento ha discurrido en términos de ejemplos. Cada uno de ellos es efectivo en sus propios términos, pero todavía son discutibles. Kant, sin embargo, afronta la cuestión de un modo tal que, si tuviera éxito, convertiría en redundante la apelación a los ejemplos. Vimos que los empiristas clásicos, como Locke, pretendían que nuestros conceptos (o ideas, para Locke) se derivan de la experiencia y trataban de establecer esta conclusión por enumeración. Kant no argumenta directamente contra ninguna de las partes de la enumeración, sino que arguye que hay algunos conceptos que no pueden derivarse de la experiencia, en la medida en que son necesarios para cualquier experiencia posible. Alguien que careciera de ellos sería incapaz de tener experiencia alguna. El argumento en este punto es extremadamente complejo, pero una idea básica es que el mundo es experimentado en la medida en que está ordenado espacial, temporal y casualmente. Por ejemplo, la experiencia de un barco desplazándose por el curso de un río es una experiencia de una cadena de sucesos que está ordenada en el espacio, en el tiempo y con relación a los vínculos causales entre sus partes. (Este ejemplo proviene de Kant, edición de 1961, B237-8.) Estos elementos de ordenación, sin embargo, no son elementos cuyo conocimiento original pudiera derivarse de experiencia alguna (como, por ejemplo, esa misma experiencia). Se trata, más bien, de que sólo podemos tener una experiencia de este tipo si poseemos previamente los conceptos que constituyen la matriz causal espacial y temporal dentro de la que se presentan los sucesos de experiencia. Por tanto, tales conceptos no son empíricos. Aunque sean discernibles en la experiencia, no pueden adquirirse extrayéndolos del contenido de ninguna experiencia. Tal extracción, si fuera posible de algún modo, sólo podría realizarla quien los posea, dado que sólo alguien así puede tener una experiencia de la que extraerlos.

Podríamos replicar a esta sugerencia que Kant no pudo distinguir entre la afirmación de que todos los conceptos son empíricos y la

afirmación de que todo el conocimiento es empírico. Los conceptos son constituyentes de las proposiciones, y nuestro conocimiento de que las proposiciones son verdaderas puede ser empírico, incluso cuando esas proposiciones tienen como constituyentes conceptos que no son empíricos. Pero, en este preciso momento, comienza a parecer discutible la distinción entre conceptos y proposiciones (si es que todavía no lo era). Esa distinción equivale, en el presente contexto, a la sugerencia de que es de algún modo posible poseer conceptos como espacio, tiempo y causalidad sin conocer la verdad de ninguna proposición de la que esos conceptos sean constituyentes. Lo que es altamente improbable. Tener el concepto de causalidad es conocer la diferencia entre sucesos que están causalmente conectados y sucesos que no lo están: entender qué es que dos sucesos están causalmente conectados. Tener el concepto de tiempo es ser capaz de colocar los sucesos en orden lineal. Encontramos, por ello, que alguien que posea el concepto de causalidad ya tiene algún conocimiento proposicional; Kant afirma, de hecho, que tal persona ya sabe que todo suceso es precedido por otro, del que se sigue de acuerdo con una regla (edición de 1961, B238-41). De un modo semejante, quien tenga el concepto de espacio conoce, de acuerdo con Kant, que el espacio es infinito, homogéneo y euclideano, y quien tenga el concepto de tiempo ya sabe que el tiempo es lineal. De modo que, para Kant, el conocimiento conceptual requiere conocimiento proposicional, por lo que las proposiciones que aparecían en nuestros ejemplos, si son conocidas de algún modo, lo son *a priori*. Sin embargo, no son analíticas, dado que constituyen conocimiento sustancial del mundo empírico, el mundo que experimentamos.

Podríamos poner objeciones a la versión kantiana del conocimiento proposicional requerido para la posesión de conceptos como espacio, tiempo y causalidad, pero es difícil discutir la afirmación de que el conocimiento conceptual requiere algún tipo de conocimiento proposicional. Por tanto, una posible objeción a Kant deberá concentrarse en su argumento de que algunos de nuestros conceptos no son empíricos.

Lo que hasta ahora hemos visto es, pues, que, si admitimos la distinción entre conocimiento *a priori* y conocimiento empírico, le será difícil al empirista mantener su afirmación de que no puede conocerse *a priori* ninguna verdad que no sea analítica. Parece que una versión más eficaz de empirismo negará la distinción entre lo empírico y lo *a priori*, manteniendo que ningún tipo de conocimiento se distingue por ser una cosa más que la otra. Ésta es la posición de Quine; nos aproxi-

maremos a ella, preguntándonos cuál es el objeto del conocimiento *a priori*. ¿Cuál es la naturaleza de esas verdades, si las hay, que pueden conocerse sin apelar a la experiencia? Kant sostuvo que cualquier verdad cognoscible *a priori* debería ser tanto universal como necesaria (edición de 1961, B3-4). Examinaremos sucesivamente cada una de esas afirmaciones.

14.4. CONOCIMIENTO *A PRIORI* Y VERDAD UNIVERSAL

¿Por qué mantuvo Kant que el conocimiento *a priori* debe ser conocimiento de una verdad universal? La razón es la de que, si el conocimiento lo es de la naturaleza de un objeto particular, deberá depender del examen de ese objeto y, por tanto, deberá ser empírico. De hecho, Kant no vio ninguna distinción entre necesidad y universalidad y supuso, por tanto, que, dado que el conocimiento *a priori* lo es de una verdad necesaria, debe serlo de una verdad universal. Concedamos por el momento que el conocimiento *a priori* lo es de verdades necesarias, y utilicemos tal concesión para argumentar que no es preciso que, por esa razón, sea también universal.

Kant no necesitaba negar la posibilidad de conocer verdades necesarias sobre objetos particulares, en la medida en que esas verdades se conozcan por inferencia a partir de verdades universales antecedentes. Así, es posible que yo sepa que el objeto que está frente a mí no puede ser, a la vez, cuadrado y no cuadrado, pero tal conocimiento se adquiere por inferencia a partir de la Ley de No Contradicción (que es universalmente verdadera). Es posible que las verdades universales y necesarias tengan verdades necesarias particulares como sus consecuencias, pero la dirección de descubrimiento —para que éste sea *a priori* y no empírico— debe ser siempre desde lo universal a lo particular. Puesto que, para que el conocimiento universal sea *a priori*, no podemos llegar a él a partir de nuestro conocimiento de los casos particulares; dado que el conocimiento particular ha de ser empírico, también lo sería el conocimiento universal derivado de él.

La debilidad del punto de vista de Kant, en este aspecto, radica en la afirmación de que todo conocimiento de verdades necesarias sobre objetos particulares ha de ser empírico, por extraerse del examen de tales objetos. Es posible que haya un modo de adquirir conocimiento *a priori* que, aunque relativo a la naturaleza de un objeto particular, se adquiera no examinándolo, sino reflexionando sobre su naturaleza.

Es importante que aclaremos lo que intentamos mostrar. Hay dos

niveles. En primer lugar, podríamos encontrar un ejemplo de verdad necesaria sobre un objeto particular que no es conocida empíricamente, aunque sí podría haberse conocido de ese modo. En segundo lugar, podríamos encontrar un ejemplo de verdad necesaria sobre un objeto particular que sea conocida, pero que no podría ser conocida empíricamente. En el primer caso, nos encontramos con el ejemplo de algo que puede ser conocido tanto empíricamente como *a priori*. En el segundo caso, tenemos un ejemplo que sólo puede ser conocido *a priori*. En los dos casos, tenemos un ejemplo de conocimiento *a priori*, dado que conocimiento *a priori* es el conocimiento que *puede* ser adquirido sin apelación a la experiencia. Es irrelevante si la misma proposición puede ser conocida de alguna otra manera. De modo que la persona que pretende que su experiencia muestra la verdad de la ley causal de que cada suceso se sigue de algún otro de acuerdo con alguna regla puede estar en lo cierto. Pero, aun así, nuestro conocimiento de esta regla es todavía *a priori*.

Kripke proporciona un ejemplo del tipo que buscamos [Kripke (1971)]. En una conferencia, le pidió a su auditorio que considerara la mesa de madera sobre la que estaba apoyado. Sus oyentes sabían, empíricamente, que se trataba de una mesa de madera. Kripke preguntó si esa misma mesa podría haber sido una mesa de hielo. Ciertamente, podría haber sucedido que hubiera habido una mesa de hielo en el lugar, en vez de una de madera, pero tal cosa no es suficiente. Lo que preguntaba era si una mesa que de hecho es de madera podría haber sido de hielo; tenemos la fuerte intuición de que ninguna mesa de hielo podría haber sido la mesa (de madera) sobre la que Kripke estaba preguntando. Podemos advertir, reflexionando sobre esa mesa, que, habiendo sido hecha de madera, no podría haber sido hecha de hielo. Éste no es un conocimiento empírico. Nosotros (o los oyentes) sabíamos empíricamente que la mesa era de madera, (p), y vimos que tal cosa quería decir que no podía haber sido hecha de hielo ($\Box q$; necesariamente, no estaba hecha de hielo). Nos encontramos, pues, con esta situación:

1. sabemos empíricamente que p,
2. sabemos que ($p \rightarrow \Box q$),
3. sabemos que $\Box q$.

Nuestro conocimiento de $\Box q$ es empírico, porque se deriva por inferencia de nuestro conocimiento empírico de que p. Pero ¿qué

sucede con nuestro conocimiento de que ($p \rightarrow\Box$ q)? ¿Es empírico o *a priori*? No parece ser empírico, porque no podemos concebir ninguna experiencia que lo verificara o lo refutara (como argumenta Kripke, ibid., pp. 52-53). De modo que se trata de un condicional conocido *a priori*, y que es una verdad necesaria.

Debemos enfrentarnos ahora a la cuestión de si la verdad necesaria $\Box(p \rightarrow\Box\ q)$, que es una verdad sobre un objeto particular, se conoce reflexionando sobre la naturaleza de un objeto particular o por subsunción como una simple instancia de una verdad universal. ¿Conocemos directamente la verdad universal de que los objetos de madera no podrían haber sido hechos de hielo, o podemos inferir la verdad universal de nuestro conocimiento del caso particular?

Creo que, al menos, el auditorio de Kripke se dio cuenta de $\Box(p \rightarrow\Box\ q)$, no reflexionando sobre las consecuencias universales de la propiedad de estar hecho de madera, sino reflexionando sobre *esa* mesa y sobre si cualquier mesa hecha de hielo podría haber sido idéntica a ella. Por supuesto, una vez advertida esa verdad con relación a esta mesa, podrían advertir que reflexiones semejantes serían igualmente efectivas en cualquier caso relevantemente similar. Lo que descubrieron *a priori* era reconocible y universalmente verdadero, y la verdad particular era, en ese sentido, un caso particular de una verdad universal. Pero no fue ésa la ruta que siguieron. Descubrieron que era una verdad necesaria reflexionando sobre la naturaleza del caso particular; su conocimiento *a priori* no era entonces universal, y no lo hubiera llegado a ser si no hubieran reflexionado más.

Lo que sugiere este ejemplo de Kripke es que, del mismo modo que una verdad universal contingente ordinaria (por ejemplo, que todos los cuervos son negros) puede conocerse por inferencia a partir de verdades particulares contingentes (sobre este u otro cuervo), una verdad universal necesaria (por ejemplo, que ninguna mesa de madera podría haber sido de hielo) puede conocerse por inferencia a partir de verdades necesarias particulares (verdades necesarias sobre mesas particulares). Pero esto no haría todavía que el conocimiento universal fuera empírico. Todavía es *a priori* porque el conocimiento particular del que se infiere es *a priori*. Kant estaba equivocado, pues, al suponer que las verdades sobre objetos particulares sólo pueden conocerse de un modo empírico. Algunas verdades necesarias sobre objetos particulares pueden conocerse *a priori*.

De hecho, podríamos intentar dar otra vuelta de tuerca al argumento de Kripke. Hasta ahora, hemos admitido que la verdad necesaria de que esta mesa, al ser de madera, no podría haber sido de hielo podía

haber sido conocida o bien directamente, o bien por inferencia a partir de la verdad universal de que ningún objeto de madera podría haber sido hecho de hielo. Pero es posible que existan algunas especiales verdades necesarias que sólo puedan descubrirse a partir de los casos particulares; en tal caso, las verdades universales inferibles de ellas no pueden reconocerse más que en la medida en que figuran en los casos particulares.

El ejemplo de Kripke puede ser de este tipo, dado que es posible que nuestro conocimiento de que ninguna mesa de madera podría haber sido hecha de hielo sólo pueda alcanzarse confrontando cuestiones sobre mesas particulares (reales o imaginarias). Puede haber otros ejemplos relativos a nuestro conocimiento de las verdades morales. Es posible que tales verdades sean necesarias, puesto que, si las acciones son mejores por ser generosas, es difícil imaginar un mundo posible en el que tal cualidad las hiciera peores (aunque pudieran ser malas por otras razones). Pero quizá sea imposible conocer la verdad necesaria de que una acción es mejor por ser generosa si no es reflexionando sobre casos particulares; es esto lo que los teóricos de la ética denominan inducción intuitiva [véase Broad (1930), p. 271].

14.5. CONOCIMIENTO *A PRIORI* Y VERDAD NECESARIA

Por ello, Kant estaba equivocado al decir que el conocimiento *a priori* debe serlo de una verdad universal. La verdad particular puede conocerse *a priori*. Pero nuestro argumento de que es así ha asumido que Kant estaba en lo cierto en su pretensión adicional de que el conocimiento *a priori* debe serlo de una verdad necesaria. Las verdades contingentes no pueden conocerse *a priori*. Y esto es así porque una verdad contingente es una que podría ser o podría haber sido falsa; una proposición contingente es una que podría ser o podría no ser verdadera, y que podría ser o podría no ser falsa. Pero, si podría ser falsa, no podemos evitar la investigación empírica si queremos descubrir si de hecho es falsa o no. La verdad contingente, entonces, sólo puede conocerse empíricamente.

Esta observación puede expresarse usando la terminología de los mundos posibles (véase 3.3). Una proposición contingente es aquella que es verdadera en algunos mundos y falsa en otros. Si deseamos saber si es verdadera en el mundo real, sólo podemos hacerlo determinando empíricamente si el mundo real es de uno u otro tipo, es decir, determinando la clase de mundos posibles en la que se incluye el mundo real.

La verdad contingente, en ese caso, sólo puede ser conocida empíricamente. Pero lo inverso no es verdad, como mostramos en la última sección. Algunas verdades necesarias sólo pueden conocerse empíricamente; por ejemplo, nuestro conocimiento de que esta mesa no podría haber sido hecha de hielo.

Lo que estamos diciendo aquí no es que, a veces, cuando *p* es necesariamente verdadera, podemos conocer empíricamente que *p* es verdad. La afirmación es, más bien, que, cuando *p* es verdad necesariamente, podemos conocer empíricamente que *p* es necesariamente verdad.

Pero el ejemplo que se ofreció era el de una verdad necesaria que se conocía empíricamente porque se conocía por inferencia a partir de dos proposiciones, una de las cuales se conocía empíricamente. ¿Podemos descubrir un ejemplo distinto de verdad necesaria que se conozca empíricamente, de un modo distinto a la inferencia? ¿Es posible tener, algunas veces, experiencia de las necesidades? La cuestión es parte de una que se planteó anteriormente, sobre el alcance del conocimiento conceptual. Por ejemplo, ¿nos es posible *ver* que algo debe ser, o no puede ser, verdadero?

El tipo de ejemplo que podríamos presentar al respecto sería, quizá, como el siguiente. Vemos que dos personas pasean por la calle, una detrás de otra. ¿Podemos ver que, para alcanzar a la otra, una de ellas tendrá que caminar más deprisa? La cuestión es, si no, ¿por qué no?

14.6. QUINE Y LA DISTINCIÓN ENTRE LO *A PRIORI* Y LO EMPÍRICO

Hasta ahora hemos decidido que, por más que todo conocimiento *a priori* lo sea de una verdad necesaria, algunas necesidades pueden conocerse empíricamente, aunque sólo sea por inferencia desde lo contingente y por medio de un condicional conocido *a priori*. Quine argumenta que no hay otra manera en la que podrían conocerse las necesidades. Ninguna necesidad puede conocerse de un modo que no sea empírico.

Esto parece implicar que Quine acepta, de algún modo, la distinción entre conocimiento empírico y *a priori*, sin asignar ningún papel ni contenido al lado del *a priori*. De hecho, Quine acepta la estrategia que en 14.3 considerábamos más prometedora para el empirismo; niega que la distinción sea mantenible en absoluto. No hay línea divisoria, ninguna razón por la que debamos obsesionarnos sobre si este o

aquel conocimiento es *a priori* o empírico. Es posible que, en este punto, haya distinciones de grado, pero todo conocimiento es más o menos empírico.

La ruta de Quine para alcanzar esta conclusión no es distinta a la ruta por la que demolió la distinción entre lo analítico y lo sintético (6.3). De hecho, el vínculo entre ambas distinciones (*a priori/empírico* y analítico/sintético) es tan íntimo que cualquier aceptación de la primera resucitaría la última. En ambos casos, el argumento parte del holismo de Quine y de la indeterminación del significado oracional. Quine argumentó, a partir de ahí, que ninguna oración de nuestra teoría es completamente inmune a la revisión. La experiencia recalcitrante causa una revisión en las oraciones de observación que forman la periferia de nuestra teoría, y, en consecuencia, una revisión en algún punto del interior. Las oraciones más cercanas a la periferia se pueden revisar más fácilmente que las oraciones que están más en el interior, pero no hay ninguna manera de discriminar por anticipado qué tipo de revisión resultará siendo más efectivo. Al ser de ese modo, incluso nuestro conocimiento más firme se ve afectado, en algún grado, por los resultados empíricos. Por el contrario, el conocimiento *a priori* no puede verse afectado por los resultados empíricos, porque no necesitamos de esos resultados para adquirirlo. Por tanto, no puede haber tal cosa como el conocimiento *a priori*. Todo el conocimiento es, en alguna medida, empírico.

La actitud de Quine hacia la distinción entre verdad necesaria y contingente resultará ahora predecible. La verdad necesaria se concibe como una verdad que no podría ser de otra manera y que, por tanto, debe ser inmune a la revisión. Si hubiera verdades necesarias, podría darse un conocimiento necesario de ellas. Pero, dado que ninguna oración es inmune a la revisión, no hay tampoco ninguna verdad necesaria.

En este punto, una matización es imprescindible. El tipo de necesidad que Quine está rechazando es necesidad lógica o conceptual, necesidad que se concibe como verdadera por la lógica, la naturaleza de nuestro esquema conceptual o los significados de nuestras palabras. Tales verdades necesarias, si se dieran, serían irreversibles y, por esa razón, Quine las rechaza. Pero existe otro tipo de necesidad, la necesidad natural, que puede aceptar, y sí acepta. Lo necesario naturalmente es aquello de lo cual la ciencia nos dice qué no podría haber sido diferente. El papel ha de arder si se le acerca un fuego y el agua ha de helarse a determinada temperatura. Las posibilidades naturales son aquellas que la ciencia permite. Este tipo de necesidad y posibili-

dad es aceptable para Quine porque las proposiciones implicadas no son más irrevisables de lo que pueda serlo la propia ciencia. Pero no hay ninguna otra noción ulterior de posibilidad lógica con la que poder decir que algunas cosas que son naturalmente imposibles (incompatibles con las leyes científicas) son, sin embargo, lógicamente posibles.

La dificultad de la posición de Quine en este punto radica en su actitud ambivalente respecto a la Ley de No Contradicción. Cuando adopta cierto talante, Quine asevera que esta ley está también técnicamente sujeta a revisión, aunque sólo sea en las circunstancias más extremas e inconcebibles. Éste es el talante de su «Dos dogmas del empirismo» («Two Dogmas of Empiricism», 1953, p. 43). En escritos posteriores, sin embargo, se muestra más dispuesto a admitir vestigios de irrevisabilidad. Argumenta [Quine (1960), pp. 57-61] que las conectivas lógicas tienen un significado determinado, lo que le permite conceder [Quine (1970), cap. 6] que las leyes lógicas son verdaderas en virtud del significado de las conectivas lógicas que en ellas aparecen. La gente que está en desacuerdo sobre las leyes de la lógica no habla en el mismo lenguaje. Lo que equivale a admitir que tales leyes son verdades analíticas irrevisables.

14.7. UN ANÁLISIS COHERENTISTA

Blanshard da un análisis un tanto diferente del conocimiento *a priori* y de la verdad necesaria.

El empirista radical se niega a aceptar la noción de necesidad natural. Para él, no hay necesidad en el mundo, sino tan sólo regularidades y nuestra respuesta a ellas. Hume sostuvo (Hume, edición de 1955, cap. 7) que nuestro sentimiento de que los objetos pesados no podrían dejar de caer, o de que una ventana de cristal está obligada a romperse cuando un ladrillo la golpea, no ha de ser explicado por la idea de que hay maneras en las que el mundo se comporta de hecho y otras en las que no puede dejar de comportarse, sino por la idea de que nuestra experiencia de regularidades perfectas origina en nosotros un hábito irresistible de tener ciertas expectativas. La necesidad natural, por tanto, es más un hecho sobre nosotros, los observadores, que sobre el mundo que observamos. La necesidad lógica, sin embargo, se admite y explica por apelación a la Ley de No Contradicción (véase Hume, edición de 1955, cap. 4.1).

Quine adopta el punto de vista opuesto. Rechaza la necesidad lógi-

ca en favor de la necesidad natural, concebida como la que está requerida por las leyes de la ciencia. Quine, por tanto, reaviva la distinción entre lo necesario y lo contingente en el seno del mundo natural, antes que como una distinción entre lo natural y la lógica.

El punto de vista de Blanshard es más bien diferente. Rechaza la vieja pretensión empirista de que toda necesidad es necesidad lógica, argumentando, más bien como hicimos en 14.3, que hay grandes cantidades de verdades sintéticas necesarias cuya necesidad parece ser un hecho sobre el mundo más bien que derivado de las complejidades de la lógica [Blanshard (1939), caps. 28-30]. Pero él extiende el dominio de la necesidad natural más lejos de lo que haría Quine, por dos razones surgidas de su coherentismo.

La primera razón es relativa a la distinción entre el conocimiento y la comprensión. Como investigadores, buscamos no sólo conocer la verdad, sino comprender, y comprender es ver por qué las cosas son como son. Al comprender, por tanto, nos movemos del reconocimiento del hecho de que algo es verdad, contingentemente verdad, al reconocimiento de que ha de ser verdad, de que es necesariamente verdad (cf. observaciones sobre la actitud coherentista respecto a la inducción en 13.5). El dominio de la necesidad natural se extiende, pues, tan lejos como pueda alcanzar la comprensión.

La segunda razón tiene un alcance incluso mayor. Para un coherentista en la tradición de Bradley, la verdad y el sistema coinciden (véase 8.2). Lo sistemático y lo coherente son lo mismo, y la coherencia se define en términos de lo mutuamente explicativo. A medida que nuestro sistema crece en coherencia, las interrelaciones entre sus partes se hacen cada vez más estrechas. El poder explicativo del sistema se hace tan grande que ya no podemos dar ningún sentido a la idea de que suceda sin más que una u otra parte sean verdaderas. En lugar de ello, comenzamos a considerar cada parte como necesaria para el todo, y, de esta manera, reconocemos que lo que anteriormente era una verdad contingente se convierte en una verdad necesaria (cf. observaciones de 8.1 sobre las diferencias entre implicación y explicación mutua). De modo que, si Blanshard está en lo cierto, la verdad contingente en último término se evapora, dejando tan sólo la verdad necesaria. Y, dado que esta necesidad ha ido aumentando progresivamente, la necesidad, como la verdad, es asunto de grado.

Esta conclusión lleva al rechazo de la distinción entre conocimiento empírico y *a priori*. Blanshard escribe [Blanshard (1939), vol. 2, p. 424]:

La percepción de que la necesidad es asunto de grado es inconsistente con la línea definida entre lo *a priori* y lo empírico que los positivistas pretenden trazar. A la luz de tales obviedades como la de que cualquier cosa roja debe ser coloreada y extensa, o la de que el naranja cae, en el espectro de los colores, entre el rojo y el amarillo, es absurdo decir que los juicios sobre un hecho empírico están completamente desprovistos de necesidad. Es también absurdo decir que un juicio determinado posee, exactamente con su significado actual, una necesidad tan absoluta como para no poder ser modificado por ninguna ampliación del conocimiento.

LECTURAS COMPLEMENTARIAS

Kant (edición de 1961, introducción) y Kripke (1971).

Putnam (1977), sobre la relación entre Quine y las distinciones analítico/sintético y *a priori/empírico*.

Stroud (1969), respecto a las ideas de Quine sobre la irrevisabilidad de la lógica. Véase también la réplica de Quine en el mismo volumen.

Pears (1966), sobre la incompatibilidad de los colores.

Las opiniones epistemológicas de Locke se encuentran en *i.* Locke (edición de 1961, vol. 4). Hume realiza una distinción análoga a la distinción analítico/sintético en Hume (edición de 1955, cap. 4.1). El empirismo de Locke se expone, no muy coherentemente, en *J.* Locke (edición de 1961, libro 4, cap. 3.25). El de Hume, en Hume (edición de 1955, cap. 12.3). Los puntos de vista de Hume sobre causalidad y necesidad natural se encuentran en Hume (edición de 1955, cap. 7).

15. ¿ES POSIBLE LA EPISTEMOLOGÍA?

15.1. HEGEL

En este último capítulo consideraremos la acusación de que la tarea de la epistemología, como ha sido concebida en este libro, no es posible. La acusación no puede defenderse con un argumento escéptico cualquiera. La mayoría de los argumentos escépticos sugieren que, por más que podamos comenzar la construcción de una teoría del conocimiento o la justificación, nunca podemos completar la tarea; al menos, si el completarla incluye demostrar que tenemos conocimiento perceptivo, científico o de cualquier otro tipo. Éste ha sido el papel del argumento del error y, hasta ahora, no hemos encontrado ninguna respuesta a él. Pero los argumentos escépticos en los que ahora estamos interesados pretenden otra cosa distinta: que la tarea epistemológica no puede ni siquiera comenzarse, y de ahí, *a fortiori*, que no puede completarse con éxito. Una manera de expresar esta observación sería decir que los argumentos del primer tipo pretenden que —aunque fuera posible construir una teoría de lo que sería el conocimiento, si tuviéramos alguno— nos encontramos con que, de hecho, no podemos pretender que conocemos nada (y, por ello, tampoco podemos pretender que sabemos que nuestra teoría es verdad). Los argumentos del segundo tipo pretenden que no podemos comenzar a construir ninguna teoría en absoluto, porque tal cosa involucraría alguna forma de circularidad viciosa o el compromiso con supuestos injustificables.

Hegel comienza la introducción a la *Fenomenología del Espíritu* considerando la acusación de que en epistemología la tarea del conocimiento es la de examinarse a sí mismo más bien que a otras cosas, y que eso es algo imposible. El conocimiento mismo no es un objeto para nosotros, sino un instrumento por medio del cual los objetos se nos aparecen. No podemos examinar el medio mismo, porque es siempre aquello a través de lo cual nos relacionamos con nuestros objetos, y si lo convertimos en un objeto dejará de ser la relación en la que estamos interesados. Esta aproximación es sugerente, pero Hegel la rechaza inmediatamente porque involucra nociones, como las de

medio o instrumento, que todavía no han sido clarificadas. Reconstruye, pues, el problema a su manera. Su primera reformulación es la de que en el estudio del conocimiento necesitamos un criterio que nos permita distinguir entre conocimiento genuino y sus sucedáneos falsificados. Pero ninguno de los criterios que pudiéramos adoptar puede haberse justificado a sí mismo al comienzo de la investigación, dado que la justificación del criterio es, precisamente, uno de los resultados que suponemos que encontraremos como conclusión. Y, sin un criterio que poder usar al comienzo, ni siquiera podemos comenzar.

Antes de proporcionar una respuesta a este problema, Hegel elabora su posición. El conocimiento es una relación entre nosotros y el objeto de nuestro conocimiento. Tal objeto, cuando lo conocemos, existe *para nosotros*. Pero el objeto de nuestro conocimiento es independiente de nosotros, y existe en *sí mismo* también. En tanto existe en sí mismo, lo denominamos verdad; en el conocimiento, nuestro fin es la aprehensión del objeto, la verdad, tal y como existe *en sí mismo*. Pero ¿cómo es posible tal cosa? ¿Cómo podemos tener un conocimiento de la verdad? Lo que sabemos, en la medida en que lo sabemos, es *para nosotros*, no en *sí mismo*. Y el criterio de lo que existe *para nosotros* radica en nosotros y viene de nosotros, mientras que aquello que juzgamos por medio de ese criterio, al existir *en sí mismo*, no necesita aceptar nuestro criterio como relevante o adecuado. El criterio constituye una imposición sobre un objeto independiente. De modo que no hay nada de lo que podamos hacer que pueda asegurar que nuestro criterio se ajusta a nuestro objeto tal y como es *en sí mismo*.

Así, el problema del criterio se convierte, en manos de Hegel, en un aspecto del problema más general de cómo es posible que nuestros esfuerzos para conocer el objeto tal y como es en sí mismo puedan tener éxito alguna vez. Y esto constituye una manera radical de expresar la dificultad ya familiar de que, independientemente de lo firmes y bien establecidas que sean nuestras creencias, siempre queda la posibilidad de que el mundo no sea como creemos. Esta dificultad, que yace a la base del argumento escéptico del error, pasa a ser el argumento de que la epistemología, más bien que no poder completarse, ni siquiera puede comenzar.

Hegel comienza su solución al problema señalando que, en epistemología, el objeto de investigación es nuestra propia aprehensión o nuestra conciencia. Pero este objeto es de tal naturaleza que la distinción entre lo que existe *para nosotros* y lo que existe en *sí mismo* no es aquí la distinción entre lo que nos es accesible a la conciencia y lo que

no nos lo es. Ambos lados de la distinción caen dentro del alcance de la conciencia. La razón es que nuestra cuestión es la de cómo estamos relacionados a nuestro objeto en la conciencia, y cuando nuestro objeto es nuestra propia conciencia está claro que no hay peligro alguno de que tal conciencia tenga una existencia *en sí misma* que, en principio, nos esté oculta y sea diferente a la conciencia tal y como existe *para nosotros*. Por tanto, en epistemología comparamos el objeto del que somos conscientes con nuestra conciencia de él, pero ello no significa que la empresa sea imposible, dado que las distinciones cruciales entre conciencia y objeto, o entre el *para nosotros* y el *en sí*, son distinciones en el seno de lo que nos es accesible, más bien que entre algo que cae dentro y algo que cae fuera del alcance de nuestra aprehensión.

Pero, aunque no haya ninguna imposibilidad general que impida la comparación entre conciencia y objeto, siempre existe la posibilidad de que, en la comparación, descubramos que el objeto tal y como es en *sí mismo* sea una cosa para la conciencia, mientras que nuestro conocimiento de ese objeto, el objeto tal y como *es para nosotros*, sea otra; y que, cuando los comparemos, no lleguen a corresponderse. Cuando sucede tal cosa, ¿qué opciones nos quedan? Una sería la de intentar alterar nuestro conocimiento para hacer que se ajustara al objeto. Pero ello no tendría el efecto deseado, por una importante razón. Nuestro conocimiento, en este caso, es el estado de conciencia que se ajusta mejor al criterio que estamos utilizando. Incluso así, no llegó a ajustarse a su objeto. Y el criterio por el que se produjo este desajuste lo encuentra la conciencia en la naturaleza del objeto, no, como se sugirió anteriormente, como algo impuesto arbitrariamente desde el exterior. De modo que no podemos esperar ganar nada «alterando» nuestro conocimiento a menos que cambiemos también el criterio; y no podemos cambiar el criterio sin cambiar el objeto. Alterar nuestro conocimiento querría decir cambiar nuestro objeto; cuando conocimiento y objeto dejan de corresponderse, ambos se vienen abajo y, con ellos, perdemos el criterio que estábamos usando para determinar si se corresponden.

Lo que sucede aquí es que, a cierto nivel de conciencia, ha surgido lo que Hegel denominaría una contradicción. Pero la actitud de Hegel ante esas contradicciones no es de desesperación. En vez de ello, supone que el resultado de esas contradicciones es que, la naturaleza de la contradicción que hemos expuesto nos puede enseñar adonde llegar desde ese punto. Cuando conocimiento y objeto no se corresponden, no nos quedamos sin nada en absoluto. Nuestros intentos de alterar el conocimiento llevan a un cambio en el objeto y, con ello, a un

nuevo par para examinar, para ver si se corresponden. Sucede así porque la negación contenida en la contradicción no es una negación *abstracta*, en la que cada parte destruye a la otra, sin dejarnos nada, sino lo que Hegel denomina una negación *determinada*, que tiene cierta forma. Y, por medio de esta forma, nos conduce desde la contradicción a un nuevo nivel, donde encontramos un nuevo objeto y una nueva investigación.

Por considerar un ejemplo ordinario (no es de Hegel): supongamos que en nuestro examen de la percepción sensorial como forma de conciencia determinamos que la percepción sensorial, tal como la concebimos, no se ajusta a su objeto; que el objeto tal y como es *en sí mismo* no se ajusta al objeto que existe *para nosotros*. Una respuesta a esto es el escepticismo sobre el conocimiento empírico. Otra es decir que, por este desajuste (la contradicción), emerge una nueva forma de conciencia, con su nuevo objeto, que requiere un nuevo examen. Y este progreso, de una forma de conciencia a otra, continuará hasta que alcancemos una forma en la que la distinción entre el objeto *para nosotros y* el objeto en *sí mismo* desaparezca por completo, porque hayamos alcanzado una conciencia cuyo objeto sólo existe en *sí mismo* en la medida en que existe *para nosotros*. Por lo menos aquí, la conciencia ya no parece obstaculizada por lo que es externo a ella; hemos alcanzado un nivel en el que la conciencia puede ser y es su propio objeto.

El progreso desde una forma de conciencia a la otra es la marca de la concepción hegeliana de la «fenomenología». Difiere en dos aspectos importantes de la perspectiva que se ha adoptado en este libro. En primer lugar, trata el argumento escéptico no como un peligro del que debamos huir, sino como una fuente de descubrimientos. Puesto que es sólo a través de argumentos escépticos como las contradicciones inherentes a un particular nivel de conciencia se nos van a revelar, es el escéptico el que muestra que, en un área determinada, nuestro conocimiento no se corresponde a nuestro objeto, y, por ello, es el escéptico quien nos conduce de una forma de conciencia a otra, hasta que alcanzamos una forma en la que conciencia y objeto coinciden. La noción crucial aquí es la de negación determinada; éste es el tipo de negación que concierne al escepticismo y que constituye el hilo conductor de la fenomenología. En segundo lugar, la epistemología es posible para Hegel, pero sólo en la medida en que adopta esta ruta, progresando de una forma de conciencia a la siguiente. Cualquier forma de epistemología que no progrese de este modo (la de Kant, por ejemplo) es destrozada por la circularidad viciosa. La apelación a esta

progresión era el único medio que vio Hegel para solucionar el problema del criterio y de superar la separación entre el *en sí mismo* y el *para nosotros*.

15.2. CHISHOLM Y EL PROBLEMA DEL CRITERIO

¿Hay alguna respuesta al problema del criterio que plantea Hegel y que no nos aleje tanto de la tradición analítica, de la que trata este libro? Pocos escritores de esta tradición han encarado este problema metodológico, pero Chisholm no sólo lo ha expuesto, sino que ha intentado proporcionar una respuesta al mismo [Chisholm (1977), cap. 7].

Chisholm comienza considerando una dificultad general. En epistemología, tendemos a comenzar con ciertas intenciones; por ejemplo, intentamos mostrar que son posibles el conocimiento matemático y el empírico. Hay, pues, ciertos tipos de conocimientos, ciertos ejemplos, que esperamos que sean validados por nuestra teoría; el que tal cosa no suceda afectará a la teoría y no a los ejemplos. Hay pretensiones de conocimiento que son más dudosas, como las aseveraciones de conocimiento religioso o moral, por ejemplo, cuyo último destino tendrá que ser decidido por la mejor de nuestras teorías. Pero nos aferraremos a algunos de los ejemplos. Y lo que, como filósofos, hacemos es extraer de los ejemplos en los que estamos de acuerdo ciertos criterios que validarán retroactivamente tales ejemplos y que nos proporcionarán una decisión sobre los ejemplos más discutibles. Ésta es la perspectiva de «sentido común» que ejemplifica la propia teoría del conocimiento de Chisholm. Aceptando que es posible el conocimiento empírico y estando persuadido por el argumento del regreso de justificaciones (4.1) de que nuestra teoría debe tener una estructura fundamentalista, se limita a formular una serie de principios epistemológicos con el efecto deseado. La justificación de los principios es, simplemente, que tiene ese efecto; y aceptamos entonces su veredicto sobre los ejemplos problemáticos.

Un punto de vista alternativo es adoptar ciertos criterios como posición de partida, dejando abierta la cuestión de qué pretensiones de conocimiento serán validadas por ellos (si es que alguna lo es). Por ejemplo, podríamos establecer como nuestro criterio la .pretensión de que todo conocimiento se derive, de algún modo, de la experiencia sensorial. Pero podría suceder que no fuéramos capaces de mostrar cómo cualquier conocimiento interesante (por ejemplo, de objetos no vistos, del pasado o de las verdades de la matemática) se deriva de la

experiencia sensorial, en cuyo caso abandonaríamos esas áreas. Así pues, según esta perspectiva, tenemos más confianza en nuestro criterio que en nuestros ejemplos y concedemos la posibilidad de que ninguno de nuestros ejemplos favoritos sobreviva a las exigencias de nuestros criterios.

Ambas perspectivas involucran supuestos para los que ellas mismas no proporcionan medios de confirmación. El punto de vista del sentido común es manifiesta expresión de un prejuicio filosófico, mientras que la perspectiva «criterial» alternativa, aunque es probable que dé algunos buenos criterios como punto de arranque, será incapaz de explicar por qué está justificada la elección de esos criterios más que la de otros. Por supuesto, es posible que la perspectiva de sentido común no deje espacio alguno para el escepticismo filosófico. Pero, en la medida en que tal cosa depende de nuestros prejuicios, difícilmente podrá reivindicarse como una ventaja.

Puesto que su propia teoría adopta la perspectiva del sentido común, no deja de ser una sorpresa la sugerencia de Chisholm de que hay una escapatoria entre los dos polos del dilema planteado en las líneas anteriores. Denomina *cognitivismo crítico* esta tercera perspectiva.

Supongamos que, en vez de fijar ciertos ejemplos (conocimiento perceptivo, conocimiento por recuerdo, etc.) e insistir en que nuestra teoría muestra que, de hecho, tenemos tales tipos de conocimiento, nos ponemos de acuerdo, por anticipado, en ciertas «fuentes» de conocimiento: percepción, memoria, razón y autoaprehensión, por ejemplo. Cuando lleguemos a la consideración de un concepto discutible, como el de conocimiento ético, podría parecer que sólo nos quedan dos opciones. O bien nos aferramos a nuestra lista original de fuentes y aseveramos que, dado que no existe ninguna más, no existe el conocimiento ético, o bien concedemos que debe haber una fuente adicional (quizá la intuición) que haga posible el conocimiento ético. Éste es el dilema original, expresado ahora en términos de las fuentes de conocimiento. Pero el cognitivismo crítico nos ofrece una tercera posibilidad. Concede que ninguna de las fuentes anteriores es una fuente directa de conocimiento ético, pero intenta mostrar que lo que tales fuentes nos proporcionan nos hace capaces de conocer determinados hechos morales. Nuestras fuentes nos proporcionan conocimiento que sirve de *signo* de la verdad ética, que *expresa* para nosotros la verdad ética y *a través del cual* podemos conocerla. Por ejemplo, el que detestemos una acción sirve para hacer que conozcamos la maldad de algo; la conducta de otra persona puede expresarnos su dolor, por lo que podemos conocer su dolor a través de su conducta.

En general, esta sugerencia puede tener sus méritos. Parece ofrecer una mayor flexibilidad a aquellos empeñados en elucidar el concepto de percepción moral o en intentar analizar nuestro conocimiento empírico de los estados mentales de otros sin introducir una fuente adicional denominada simpatía. Así formulado, esto no es tanto una teoría como un prometedor anticipo de teoría. Para ir más allá, nos gustaría saber en qué consiste que una cosa exprese otra, en qué consiste ver una cosa como expresión de otra, y si, en diferentes áreas, por ejemplo, en ética y en las otras mentes, hay distintas formas de expresión.

Pero el problema es, en realidad, por qué tiene que pensar Chisholm que el cognitivismo crítico es una tercera posibilidad para escapar al dilema que se planteaba originalmente. El cognitivismo crítico parece especialmente preocupado con la manera de aproximarnos a los casos más discutibles, particularmente de conocimiento moral y religioso (quizá también de conocimiento del futuro). Pero el problema del criterio está vinculado, en realidad, a nuestra actitud no respecto a los casos discutibles, sino respecto a los indiscutibles. El problema era que la selección inicial de cuatro fuentes de conocimiento no era nada más que un prejuicio filosófico, y nada en el cognitivismo crítico sirve de respuesta a esta queja.

Hegel pensaría que una epistemología que comenzara con la aceptación de las cuatro fuentes sería algo ridículo (un cognitivismo acrítico), tanto por aceptar como punto de partida algo sin ninguna posibilidad de ser validado por el procedimiento, como porque, si la distinción entre las diferentes fuentes de conocimiento se toma seriamente como el punto de partida de una teoría de la mente y el conocimiento, se vendría abajo por sus propias contradicciones. Ninguna de las fuentes, aisladamente consideradas, es capaz de darnos conocimiento. La conciencia no puede corresponderse, en ninguno de los casos, a su objeto; cada nivel de conciencia crea sus propios criterios, y el análisis muestra que no pueden satisfacer los criterios creados de ese modo. De modo que, si el problema de Hegel es real, la única conclusión posible es la de que Chisholm no lo ha respondido.

15.3. QUINE Y LA INEXISTENCIA DE LA FILOSOFÍA PRIMERA

Hegel y Chisholm consideran la acusación de que la epistemología debe o bien asumir un criterio que no está en situación de justificar, o bien considerar como dados ciertos criterios cuya elección es

un mero asunto de prejuicios. Quine también se enfrenta a este problema. El holismo que resulta de la aceptación, por su parte, de la tesis de Duhem tiene, como una de sus consecuencias, el abandono de la distinción analítico/sintético, es decir, la distinción entre las oraciones que son verdad simplemente en función de su significado y aquellas cuya verdad depende, al menos hasta cierto punto, de cómo sea el mundo (véase 6.3). Aunque, para Quine, haya una diferencia de grado entre las oraciones a cuya verdad estamos firmemente comprometidos y aquellas para cuyo abandono se nos puede persuadir más fácilmente, no hay oraciones que sean completamente irrevisables. En cierta manera, todas las oraciones han de considerarse sintéticas, aunque algunas sean más sintéticas que otras. Esta versión de holismo nos fuerza a abandonar cualquier esperanza de *filosofía primera*, un sistema filosófico que se mantenga separado de (que se justifique con independencia de, y que pueda evaluar a) las afirmaciones de las ciencias especiales como la física, o, en términos más mundanos, la percepción sensible. La filosofía, y en particular la epistemología, forma un continuo con la ciencia natural, o es incluso una parte de ella. No es una investigación peculiar de nuestros conceptos, ni una investigación independiente de los significados de términos cruciales como «saber» o «justificar». Si existiera de hecho una filosofía primera, es posible que fuera éste su objeto específico. Pero la filosofía sólo se distingue de otras formas de investigación humana por su generalidad, se aferra a cuestiones que son más generales y amplias que las que investigan las ciencias especiales de la física y la psicología.

De acuerdo con esta perspectiva quineana, la filosofía es el estudio de la ciencia desde el mismo interior de la ciencia. Pero esto parece plantear problemas de circularidad. Al estudiar la ciencia desde el interior de la ciencia, el filósofo no es capaz de cuestionar de un solo golpe la totalidad de la ciencia; ha de asumir, más bien, la validez general de los procedimientos y resultados científicos para poder encontrar razones en el interior de la ciencia que le permitan cuestionar, aceptar, rechazar o reemplazar aspectos particulares.

Ésta es la razón por la que Quine se muestra tan complacido con la parábola de Neurath del marinero obligado a reconstruir su barco mientras ha de mantenerse a flote sobre él. Debemos mantener el barco de la ciencia intacto, en líneas generales, mientras lo examinamos y reparamos sus partes más defectuosas. No podemos amarrarlo a un puerto seguro y salimos de él, ni podemos suponer que el descubrimiento de contradicciones en el seno de la ciencia haya de permitirnos

elevarnos por encima de ella y abandonar el barco a bordo de una especie de helicóptero hegeliano.

¿Qué es lo que persuade a Quine de que esta concepción de las relaciones entre filosofía y ciencia le capacita para evitar la acusación hegeliana de circularidad? Parece como si, al abandonar la filosofía primera, la acusación reaparece de nuevo en una forma mucho más dañina.

Una de las sugerencias de Quine es la de que el problema de la circularidad sólo surge en el seno de una tradición filosófica obsesionada por la búsqueda de la certeza y por la posibilidad de deducir la ciencia a partir de los datos de los sentidos [Quine (1969), pp. 83-84]. Pero esta sugerencia no basta. Es verdad que, si intentáramos deducir la ciencia de los datos de los sentidos, caeríamos en un círculo vicioso si apeláramos a la ciencia para nuestros propósitos. Pero es falaz suponer que todo irá bien en cuanto renunciemos a la esperanza de semejante deducción.

La sugerencia de Quine en este punto sólo comienza a tener sentido en el seno de su punto de vista más general. Dado que, en ausencia de la filosofía primera, no nos queda otra alternativa que la de examinar la ciencia desde dentro, no hay ningún peligro de que la filosofía deba adoptar o imponer un criterio desde el exterior. Los criterios a usar son los criterios de la ciencia, lo que no involucra ninguna circularidad o supuesto previo, sino el mero reconocimiento de cuál es la tarea propia de la epistemología. En segundo lugar, las únicas dudas escépticas que son posibles son aquellas que, más que intentar criticar la ciencia desde un arbitrario patrón de «racionalidad», se derivan de la misma ciencia. La cuestión que nos planteemos no debe ser «¿qué es lo que hace que nuestras creencias científicas cuenten como conocimiento?», dado que tal cuestión nos lleva a suponer que ninguna respuesta a ella puede apelar, sin circularidad a los resultados de la ciencia. En vez de eso, nos debemos preguntar: «Si nuestra ciencia fuera verdadera, ¿cómo lo podríamos saber?» En este caso, la cuestión epistemológica se plantea dentro del alcance de un condicional y, dado que la cuestión supone la verdad de los descubrimientos científicos contemporáneos, la respuesta lo puede hacer también. La epistemología aparece aquí en el interior de la ciencia. De modo que los peligros que Hegel señala sólo surgen, para Quine, en aquellos sistemas que separan ciencia y filosofía. Una vez que, como holistas, dejemos a un lado semejante separación y nos hagamos menos ambiciosos, ya no podrá haber ninguna objeción metodológica general a la práctica de la epistemología.

15.4. EPISTEMOLOGÍA NATURALIZADA

¿En qué se convierte la filosofía, una vez que se concibe como parte de la ciencia, más que como un estudio independiente de ella? ¿Vamos a ser capaces de preguntar (y de esperar una respuesta a) las mismas cuestiones que preguntábamos anteriormente? Y, ahora, ¿qué actitud deberemos adoptar con el escepticismo? ¿Queda algún lugar para el escéptico?

La epistemología tradicional estudiaba la relación entre los datos y las creencias, entre la evidencia y la teoría. Intentaba mostrar cómo nuestras creencias (por ejemplo, la creencia en un mundo externo) están justificadas por los datos de los que surgen; cómo nuestras teorías científicas están justificadas por la evidencia en la que descansan. Este estudio ¿ha de ser abandonado y reemplazado, o puede continuarse en el seno de una nueva perspectiva? Quine parece vacilar entre estas dos alternativas. A veces, sugiere que las viejas cuestiones huelen a filosofía primera y que, de cualquier modo, el intento de descubrir una relación entre evidencia y una teoría que justificara la teoría se ha probado infructuoso. En ese caso, pregunta por qué no limitarnos a estudiar simplemente cómo nos comportamos, de hecho, cuando nos movemos desde nuestros datos a la formación de la creencia. Este estudio fáctico, claramente dentro de los límites de la psicología, es lo que denomina *epistemología naturalizada*. Deja a un lado cuestiones relativas a la justificación y considera sólo las cuestiones genéticas, o causales. Dejamos de preocuparnos sobre la brecha entre evidencia y teoría, y estudiamos las relaciones causales entre las dos.

Quine tiene alguna sugerencia sobre el modo en que podría desarrollarse este tipo de investigación. Encuentra un modelo de la práctica de la construcción de teorías en la práctica del aprendizaje del lenguaje. Las oraciones de observación son básicas en cada caso; son la evidencia sobre la que descansan nuestras teorías y el lugar en el que el lenguaje confronta la realidad de un modo lo suficientemente directo como para que las oraciones individuales puedan aprenderse (véase 7.2). De modo que un sucedáneo empírico para el estudio de la relación entre evidencia y teoría es el estudio de las maneras en las que los que aprenden el lenguaje se mueven, de hecho, desde la comprensión de las oraciones simples de observación a la comprensión de las oraciones más complejas (que expresan disposiciones y tendencias o las consecuencias de condiciones no satisfechas) de las que se construyen las teorías.

Lo que Quine sugiere aquí es que la epistemología naturalizada no

implica un cambio de tema, sino que ofrece, por el contrario, una manera nueva de estudiar un tema viejo. El viejo problema era el de la brecha entre «*input* escaso» y el «*output* torrencial». Pero esta brecha puede estudiarse de dos maneras distintas, o bien estudiando la relación entre oraciones de observación y oraciones teoréticas, como mencionábamos anteriormente, o bien, más directamente, estudiando la relación entre el *input* físico recibido por el sujeto humano —las modificaciones en la retina, por ejemplo, constituyen la información recibida por el ojo— y las creencias del sujeto causadas por él; siendo estudiadas esas creencias de un modo fisicalista, es decir, estudiando la neurofisiologia de la actividad cerebral que las constituye. Este último punto de vista es, quizá, el más característico de la epistemología naturalizada, y Quine sostiene (Quine, 1969, p. 83) que nos vemos obligados a adoptarlo,

> por razones que, de algún modo, son las mismas que han impulsado siempre la epistemología, para poder ver cómo se relaciona la evidencia con la teoría, y de qué manera la teoría que uno tenga sobre la naturaleza trasciende cualquier evidencia accesible.

En lo que respecta a los argumentos escépticos, ya hemos visto que Quine está dispuesto a admitir la cuestión epistemológica general: «Si nuestra ciencia fuera verdadera, ¿cómo lo sabríamos?» Al escéptico sí se le concede un papel en el seno de este planteamiento. Deberá encontrar razones interiores a la ciencia para cuestionar la posibilidad de conocer la verdad científica; nuestra ciencia deberá mostrarse a sí misma como incognoscible. Según Quine, tal cosa no es imposible, aunque sí sea muy improbable. Hay dos estrategias escépticas habituales que están prohibidas de antemano por el requisito de que el escéptico trabaje desde el interior de la ciencia.

La primera la constituye cualquier versión del argumento del error que comienza con la afirmación de que es lógicamente posible, en cualquier momento y en cualquier circunstancia, que las propias creencias sean falsas. Lo que quiere Quine es hundir en el lodo cualquier argumento de este tipo, negándose a aceptar la noción de posibilidad lógica que usa. La única noción de posibilidad que está dispuesto a admitir es la de posibilidad física, aquello que nuestra ciencia acepta como posible. Permitir otro tipo de posibilidad, que fuera independiente de los resultados de la física, sería volver a generar la distinción analítico/sintético. Y no es físicamente posible que, en cualquier momento y en cualquier circunstancia, la creencia actual de uno sea falsa. Si el único espacio lógico para la falsedad fuera el espacio lógi-

co, no habría espacio en absoluto. De modo que este argumento escéptico general no puede funcionar.

De un modo semejante, el escéptico podría intentar argumentar que, por todo lo que sabemos, la realidad podría ser completamente diferente a como creemos que es; no es necesario que el mundo confirme nuestra teoría, ni que el objeto tal y como es *para nosotros* se corresponda al objeto tal y como existe *en sí mismo*. Según Quine, todas estas afirmaciones descansan en el supuesto de que hay un objeto, el mundo, que está separado de nuestra teoría y que proporciona un criterio por el que nuestra teoría puede estar determinada como falsa (por supuesto, no por nosotros, sino como una cuestión de hecho). Pero, según el análisis de Quine de las relaciones entre ciencia y epistemología, ese supuesto carece de sentido. El único criterio de realidad es el que nos proporciona la ciencia, la única realidad es la que describe la ciencia. Así que, de nuevo, no hay peligro alguno de que nuestro criterio deje de ajustarse al objeto, dado que la ciencia proporciona, a la vez, el criterio y el objeto: una situación que debería resultar, hasta cierto punto, atractiva para Hegel.

Estas dos perspectivas escépticas están, pues, prohibidas, pero cualquier argumento escéptico que use la ciencia para refutar la ciencia es, al menos metodológicamente, aceptable. Y el mismo Quine nos da uno. Ya que considera que un descubrimiento de la ciencia es que recibimos un «*input* escaso» del que generamos un «*output* torrencial». Y, ciertamente, este contraste entre lo escaso y lo torrencial es todo lo que necesita el escéptico para elaborar su argumento desde el interior de la ciencia contra la posibilidad de un conocimiento científico o teorético. Puesto que, si el contraste es tan amplio, ¿cómo es posible que haya en el *input* lo suficiente como para justificar el *output* que damos como respuesta a él?

Creo que la respuesta que daría Quine es la que repite en contextos similares [Quine (1981), p. 475]: que la reacción del escéptico es exagerada. En lugar de saltar precipitadamente a las enormes conclusiones escépticas, debemos esperar a ver en qué se convierte el estudio naturalista de la relación entre *input* y *output*. Nos puede parecer por anticipado que el *input* es desproporcionadamente pequeño para poder fundamentar el abundantísimo *output*, pero quizás la psicología empírica pueda volver a estabilizar el fiel de la balanza. De modo que el contraste entre *input* y *output* no es, todavía, un resultado de la ciencia: puede serlo o puede no serlo.

Parece que, en este punto, hay una confusión. Quine está suponiendo que la cuestión de si hay una brecha desproporcionada entre

input y *output* es empírica, y que puede responderse por un estudio naturalista de las relaciones causales entre *input*, concebido como un estímulo sensorial, y *output*, concebido como los estadios fisiológicos del cerebro que son los correlatos físicos de las creencias. Pero, desde ese punto de vista, no hay ningún contraste entre lo escaso y lo torrencial. El *input* es (conjuntamente con otras cosas) suficiente para causar los estados cerebrales que son sus efectos, y, al seguir esta historia causal, ya no estamos investigando una brecha entre el *input* y el *output*, si tal brecha se concibe análoga a la brecha entre evidencia y teoría. El contraste entre un *input* escaso y un *output* torrencial, como la brecha entre evidencia y teoría, no es un asunto *causal*, sino *inferential*.. Pertenece a lo que Sellars denomina «el espacio lógico de las razones» [cf. Sellars (1963), cap. 5.1]. La evidencia no se concibe como causalmente insuficiente para fundamentar la teoría; es insuficiente (si lo es realmente) en el sentido de no proporcionar razón suficiente para (o de no justificar) la teoría. Por ello, Quine se enfrenta a un dilema. O bien prohíbe las cuestiones inferenciales como impropias de la epistemología naturalizada, o las acepta y no puede proporcionar un método para responderlas. [Esta objeción está bien formulada por Stroud(1984), cap. 6].

Pero ¿qué nos podría justificar para aceptar la primera alternativa del dilema? Es posible que Quine afirmara, o admitiera, que, al estudiar las relaciones causales entre el *input* y lo que, a veces y un tanto tendenciosamente, denomina información, hemos abandonado el interés tradicional de los epistemólogos en las cuestiones evidenciales. Desde un punto de vista naturalista, la brecha crucial ya ha dejado de existir. Pero ¿cómo podría justificar, desde su posición, semejante renuncia? El mero hecho de que, para él, la epistemología se haya naturalizado, o de que sea ahora parte de la ciencia antes que una corte suprema, no significa que las cuestiones sobre la justificación no puedan juzgarse de ningún modo. La misma ciencia no es completamente naturalista. Contiene sus mismos criterios evaluadores, y tales criterios pueden usarse desde el interior de la ciencia para afrontar cuestiones evaluativas como las de la justificación. Parece, pues, que la epistemología naturalizada no proporciona ninguna respuesta al escéptico; ni siquiera un método por el que pudiera encontrarse tal respuesta.

La misma distinción que hemos venido usando entre cuestiones evidenciales y cuestiones causales podría ser cuestionada. ¿Por qué las mismas investigaciones causales no pueden ser investigaciones sobre la justificación? Después de todo, entre los posibles análisis de la jus-

tificación que considerábamos anteriormente, una posibilidad relevante era la teoría causal de la justificación, que sostiene que las creencias están justificadas si y sólo si están causadas de cierta manera. Es posible que Quine quisiera usar este resultado para mostrar cómo un interés en asuntos causales puede dotar de sentido, y mantener la esperanza de responder a las cuestiones sobre justificación. Pero, si ésta es una respuesta al escéptico, es sólo otra forma de externalismo en teoría de la justificación. No hemos visto, hasta ahora, razón alguna por la que debamos aceptar respuestas externalistas; es el internalismo el que tiene el respaldo de nuestras intuiciones (9.2-3). Pero, sea como sea, la adopción por parte de Quine del externalismo sería una actitud independiente, que no podría justificarse por ninguna de las tesis distintivas que se han discutido aquí. Ni la ausencia de la filosofía primera ni la naturalización de la epistemología tienen como resultado un argumento independiente en favor del externalismo. De modo que, si lo que buscamos es una respuesta internalista a las cuestiones epistemológicas, Quine no nos la da, ni tampoco ningún sucedáneo.

15.5. CONCLUSIÓN

Al comienzo de este capítulo distinguíamos entre argumentos a favor de que la tarea epistemológica no puede ni siquiera comenzar de los argumentos que defendían que nunca podrá ser completada con éxito. (Algunos argumentos eran de los dos tipos.) Hemos visto que Hegel proporciona una concepción de la epistemología como la progresión de un estado de conciencia a otros estados más elevados, bajo la que se pueden eludir los argumentos en favor de la imposibilidad de la epistemología. La idea de Chisholm del cognitivismo crítico no proporciona ninguna respuesta en absoluto, pero el rechazo de Quine de la filosofía primera tenía como resultado una perspectiva no-hegeliana que escapaba a la acusación de circularidad viciosa. La diferencia entre Hegel y Quine es, no obstante, que la perspectiva de Hegel contiene la promesa de una respuesta al escéptico, entendido ahora como alguien que argumenta que la tarea epistemológica nunca se completará con éxito. Las opiniones de Quine, por el contrario, no parecen ofrecer ninguna estrategia para refutar los argumentos escépticos que aparecerán normalmente dentro de los límites que él mismo acepta, es decir, argumentos escépticos desde el interior de la ciencia. De modo que todavía nos queda un argumento escéptico del que no sabemos cómo escapar si no seguimos la estrategia de Hegel.

Esto me obliga a ofrecer mis propias sugerencias. Una idea atractiva sobre cómo responder al escéptico, hecha, por ejemplo, por Stroud (1984, cap. 7), es que debemos encontrar algún medio de evitar que el escéptico generalice a partir de los ejemplos que él mismo escoge. Es decir, podríamos admitir que no sabemos que no somos cerebros en una cubeta y, sin embargo, seguir confiando en que evitaremos que se nos obligue a admitir que tampoco sabemos otras cosas. Ésta era la estrategia de Nozick, pero ya vimos en el capítulo 3 que no tenía éxito. Nos deja con la aseveración contraintuitiva de que podemos saber qué haremos mañana sin poder saber si existirá un mañana. Pero hay otro sentido en el que el escéptico está generalizando a partir de ejemplos aceptados, sobre el que quizás podríamos fijar nuestra atención. Este sentido puede verse claramente en las versiones habituales del argumento de la ilusión, y está también en el movimiento crucial del argumento del error.

El argumento del error sostiene que, si nuestros estados cognitivos actuales son, en la medida en que los podemos discriminar, relevantemente indiscernibles de otros que no eran conocimiento, no podemos aseverar que sabemos. Y no se trata sólo de que no podamos aseverar que sabemos; no sabemos porque nuestro estado actual es relevantemente indistinguible de (semejante en todos los aspectos relevantes) otro en el que no sabemos.

Este argumento descansa en un análogo epistemológico del principio de universalizabilidad familiar en teoría ética (véase 1.2). Según mi punto de vista, tal principio es un error, y mostrar tal cosa me parece lo importante en la lucha en contra del escéptico. (Todo el mundo dice piadosamente que su objetivo es aprender del escéptico, pero la práctica sólo se ajusta raramente a este deseo.)

El principio de universalizabilidad es un error en ética porque pasa por alto la capacidad que puedan tener otras propiedades en un caso distinto para refutar lo que han sido razones suficientes para un juicio moral en casos anteriores, sin que tengamos que reconsiderar de nuevo el caso anterior ni revisar nuestro juicio al respecto [véase Dancy (1981)]. A causa de esta capacidad, nunca podemos ser conducidos de un caso a otro por la universalizabilidad de los juicios morales. El hecho de que este caso sea indistinguible de otro en todas las características relevantes para el valor moral del primero no asegura que el segundo no posea otras características moralmente relevantes, por lo que ninguna posible evaluación del segundo puede alterar la opinión que nos hicimos del primero, a menos que decidamos que los dos casos son semejantes en todos los aspectos relevantes (en todos los

aspectos relevantes para cualquiera de ambos). Pero alguien que piense que son moralmente diferentes no tomará esa decisión, por lo que no se verá atrapado por la universalizabilidad. Su posición es consistente, en la medida en que mantenga que existe una diferencia moral relevante, aunque él no pueda señalarla todavía.

La analogía con la epistemología sugiere que no es posible mostrar que no sabemos ahora por la mera constatación de que no podemos señalar una diferencia relevante entre el caso actual y otro en el que no sabemos. En la medida en que aseveremos (como lo haremos) que hay una diferencia relevante, nuestra incapacidad de señalarla no es una prueba de que no sepamos, ni siquiera de que estemos equivocados (o que seamos inconsistentes) al afirmar que sabemos.

Sin embargo, esta conclusión no es una forma de externalismo. Alguien que sepa ahora, a pesar de no poder señalar una diferencia relevante entre la situación actual y otra en la que no sabe, todavía está, o puede estar, en posesión de los factores en virtud de los cuales su estado cognitivo es de conocimiento. Ello es así porque el hecho de que haya una diferencia relevante (y la diferencia relevante que hay) no es uno de los rasgos en virtud del cual ahora sí sabe. Esto le *permite* saber ahora cuando no sabía antes, pero las propiedades en virtud de las cuales ahora sabe no incluyen este *permiso*. Las propiedades en virtud de las cuales sabe ahora son propiedades más ordinarias sobre el caso actual, no remotas relaciones entre este caso y otros. Y estas propiedades más ordinarias son propiedades que él (probablemente) puede señalar. Por ejemplo, sé que hoy es miércoles aunque no pueda decir que no he cometido, o no pueda cometer, errores en situaciones relevantemente similares. Es posible que no pueda distinguir mi situación de otras reales o posibles, pero sí puedo decir cómo sé que hoy es miércoles, y esto es todo lo que necesita el internalismo en teoría de la justificación.

Hay una semejanza considerable entre el argumento que ofrezco aquí y la respuesta externalista (en el otro sentido de externalismo, externalismo en filosofía de la mente) al argumento de la ilusión (11.4); el argumento descansa también en una versión del principio de universalizabilidad. Cualesquiera que sean sus méritos, no puedo aseverar que mi argumento sea completamente firme. Lo que quiere decir que el título que había pensado para este libro era correcto por dos razones. Dejé a un lado ese título, *Pies de barro*, por no ser demasiado claro. Pero expresaba el sentido en el que el coherentismo mantiene que podemos tener conocimiento empírico sin una base sólida en la que apoyarnos: sin fundamentos. Expresa también el hecho de que el

escepticismo puede continuar siendo más perdurable, más seductor y más firme que cualquiera de las respuestas hasta ahora encontradas.

LECTURAS COMPLEMENTARIAS

Hegel (edición de 1931, Introducción), Chisholm (1977, cap. 7), Quine (1969, cap. 3, y 1973, pp. 1-20) y Stroud (1984, cap. 6).

Taylor (1972) argumenta que el argumento de Wittgenstein sobre el lenguaje privado es un buen ejemplo de lo que Hegel entiende por negación determinada.

La noción de Chisholm de cognitivismo crítico es similar a una idea apuntada por McDowell (1982).

Rorty (1980) es un intento más reciente de mostrar la imposibilidad de la epistemología para el que no he encontrado espacio. (Hay que tener cuidado con lo que denomina «fundamentalismo».) Rorty sugiere la necesidad de una concepción diferente de la filosofía, una nueva perspectiva.

Este libro está enraizado en la tradición que Rorty critica. Para trabajar en el seno de la tradición continental, podría comenzarse con las siguientes fuentes, primarias y secundarias: Habermas (1972, Apéndice y cap. 1); sobre Habermas, McCarthy (1978) y Ottmann (1982); Foucault (1972); sobre Foucault, Taylor (1984) y Coussins y Hussain (1984, 1); por último, Bernstein (1983) ofrece un análisis general.

REFERENCIAS BIBLIOGRÁFICAS

ALBRITTON, R. (1959): «On Wittgenstein's use of the term "criterion"», *Journal of Philosophy*, 56, 845-857. Reimpreso en Pitcher (1966).

ALSTON, W. P. (1971): «Varieties of privileged access», *American Philosophical Quarterly*, 8, 223-241. Reimpreso en Chisholm y Swartz (1973).

— (1976): «Two types of foundationalism», *Journal of Philosophy*, 85, 165-185.

ARISTÓTELES (1962): *De Interpretation*, editado con las *Categories* por J. L. Ackrill, Clarendon Press, Oxford. [Edición en castellano: *Peri Hermeneias. De Interpretatione*, Cuadernos Teorema, Departamento de Lògica, Universidad de Valencia, 1980.]

ARMSTRONG, D. M. (1961): *Perception and the Physical World*, Routledge and Kegan Paul, Londres. [Edición en castellano: *La percepción y el mundo fisica*, traducción de P. García Ferrerò, Tecnos, Madrid, 1966.)

— (1968): *A Materialist Theory of the Mind*, Routledge and Kegan Paul, Londres.

— (1973): *Belief, Truth and Knowledge*, Cambridge University Press, Londres.

AUSTIN, J. L. (1962): *Sense and Sensibilia*, Clarendon Press, Oxford. [Edición en castellano: *Sentido y percepción*, traducción de A. García Suárez, Tecnos, Madrid, 1981.]

AYER, A. J. (1946): *Language, Truth and Logic*, 2.ª ed., Gollancz, Londres. [Edición en castellano: *Lenguaje, verdad y lógica*, traducción de M. Suárez, Martínez Roca, Barcelona, 1971.]

— (1950): «Basic propositions», en M. Black (ed.), *Philosophical Analysis*, Prentice-Hall, Englewood Cliffs, NJ; pp. 60-74. Reimpreso en Ayer (1965) y en Chisholm y Swartz (l973).

— (1954): «Can there be a private language?», *Proceedings of the Aristotelian Society*, supp. vol. 28, 63-76. Reimpreso en Pitcher (1966).

— (1965): *Philosophical Essays*, Macmillan, Londres. [Edición en castellano: *Ensayos filosóficos*, Ariel, Barcelona, 1979.]

— (1969): *The Foundations of Empirical Knowledge*, Macmillan, Londres.

— (1976): *The Central Questions of Philosophy*, Penguin Books, Londres. [Edición en castellano: *Los problemas centrales de la filosofía*, traducción de R. Fernández González, 2.ª ed., Alianza, Madrid, 1984.]

BAKER, G., y HACKER, P. M. S. (1984): *Scepticism, Rules and Language*, Blackwell, Oxford.

BARKER, S. F., y ACHINSTEIN, P. (1960): «On the new *riddle of induction*», *Philosophical Review*, 69, 511-522. Reimpreso en Nidditch (1968).

BERKELEY, G. (1954): *A New Theory of Vision and Other Writings*, ed. A. D. Lindsay, Everyman Library, Dent/Dutton), Londres; contiene también *Principles of Human Knowledge* y *Three Dialogues;* 1.ª ed., 1709, 1710 y 1713, respectivamente. [Ediciones en castellano: *Tratado sobre los principios del conocimiento humano*, traducción de C. Cogolludo Mansilla, Gredos, Madrid, 1990; *Tres diálogos entre*

Hilas y Filonus en oposición a escépticos, traducción de A. Masegosa, 2.ª ed., Aguilar, Madrid, 1956.]

BERLÍN, I. (1939): «Verification», *Proceedings of the Aristotelian Society*, 1938-1939, 39, 225-248. Reimpreso en G. Parkinson (ed.), *The Theory of Meaning*, Oxford University Press, Oxford, pp. 15-34.

BERNSTEIN, R. (1983): *Beyond Objectivism and Relativism*, Blackwell, Oxford.

BLACK, M. (1954): *Problems of Analysis*, Routledge, Londres.

— (1958): «Self-supporting inductive arguments», *Journal of Philosophy*, 55, 718-725. Reimpreso en Swinburne (1974), y Nidditch (1968).

BLANSHARD, B. (1939): *The Nature of Thought*, Allen and Unwin, Londres.

BONJOUR, L. (1980): «Externalist theories of empirical knowledge», en French *et al* (1980), pp. 53-73.

BRADLEY, F. H. (1914): *Essays on Truth and Reality*, Oxford University Press, Oxford.

BRANDT, R. (1955): «The epistemological status of memory beliefs», *Philosophical Review*, 64, 78-95. Reimpreso en Chisholm y Swartz (1973).

BROAD, C. D. (1930): *Five Types of Ethical Theory*, Routledge and Kegan Paul, Londres.

CAHN, S. M. (1967): *Fate, Logic and Time*, Yale University Press, New Haven, Conn.

CARNAP, R. (1967): *The Logical Structure of the World*, Routledge and Kegan Paul, Londres.

CORNMAN, J. W. (1975): *Perception, Common Sense and Science*, Yale University Press, New Haven, Conn.

— (1977): «Foundational versus non-foundational theories of empirical justification», *American Philosophical Quarterly*, 14, 287-297. Reimpreso en Pappas y Swain (1978).

COUSINS, M., y HUSSAIN, A. (1948): *Michel Foucault*, Macmillan, Londres.

CHISHOLM, R. M. (1948): «The problem of empiricism», *Journal of Philosophy*, 45, 512-517. Reimpreso en Swartz (1965).

— (1977): *Theory of Knowledge*, 2.ª ed., Prentice-Hall, Englewood Cliffs, NJ. [Edición en castellano: *Teoría del conocimiento*, traducción de V. Peris Minqueza, Tecnos, Madrid, 1982.]

CHISHOLM, R. M., y SWARTZ, R. J. (eds.) (1973): *Empirical Knowledge*, Prentice-Hall, Englewood Cliffs.

DANCY, J. P. (1981): «On moral properties», *Mind*, 90, 367-385.

— (1984a): «On coherence theories of justification: can an empiricist be a coherentist?», *American Philosophical Quarterly*, 21, 359-365.

— (1984b): «On the tracks of the sceptic», *Analysis*, 44, 121-126.

DAVIDSON, D. (1974): «Belief and the basis of meaning», *Synthese*, 27, 309-323. Reimpreso en Davidson (1984).

— (1984): *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford. [Edición en castellano: *De la verdad y de la interpretación*, traducción de G. Filippi, Gedisa, Barcelona, 1989.]

DENNETT, D. (1978): *Brainstorms*, Bradford Books, Montgomery, Vt.

DESCARTES, R. (1955): *The Philosophical Works of Descartes*, 2 vols, edición y traducción de E. S. Haldane y G. R. T. Ross, Cambridge University Press, Cambridge. [Las dos obras más conocidas de Descartes tienen distintas traducciones al castellano. Cabe mencionar las ediciones de *Meditaciones metafísicas con objeciones y respuestas*, traducción de V. Peña García, Alfaguara, Madrid, 1977; *Discurso del método*, 2.ª ed., Tecnos, Madrid, 1990.]

DRETSKE, F. (1971): «Conclusive reasons», *Australasian Journal of Philosophy*, 49, 1-22. Reimpreso en Pappas y Swain (1978).

DUMMETT, M. A. E. (1978): *Truth and Other Enigmas*, Duckworth, Londres.

EDWARDS, P. (1949): «Rusell's doubts about induction», *Mind*, 68, 141-163. Reimpreso en Swinburne (1974).

EWING, A. C. (1934): *Idealism: A Critical Survey*, Methuen, Londres.

FEIGL, H., y SELLARS, W. (eds.) (1949): *Readings in Philosophical Analysis*, Appleton-Century-Crofts, Nueva York.

FIRTH, R. (1964): «Coherence, certainty and epistemic priority», *Journal of Philosophy*, 61, 545-557. Reimpreso en Chisholm y Swartz (1973).

FOLLESDAL, D. (1975): «Meaning and experience», en Guttenplan (1975), pp. 25-44. FLEW, A. G. N. (1954): «Can an effect precede its cause?», *Proceedings of the Aristotelian Society*, sup. vol. 28, 45-62.

FOUCAULT, M. (1972): *The Archaelogy of Knowledge*, traducción de A. M. Sheridan, Tavistock, Londres.

FRENCH, P. A.; UEHLING, T. E., y WETTSTEIN, H. K. (eds.) (1980): *Midwest Studies in Philosophy*, vol. 5, *Studies in Epistemology*, University of Minnesota Press, Minneapolis Minn.

GARRET, B. J. (1983): «Nozick on knowledge», *Analysis*, 43, 181-184.

GETTIER, E. L. (1963): «Is justified true belief knowledge?», *Analysis*, 23, 121-123. Reimpreso en Phillips Griffiths (1967).

GOLDMAN, A. I. (1967): «A causal theory of knowing», *Journal of Philosophy*, 64, 355-372. Reimpreso en Pappas y Swain (1978).

— (1976): «Discrimination and perceptivo knowledge», *Journal of Philosophy*, 73, 771-799. Reimpreso en Pappas y Swain (1978).

— (1979): «What is justified belief?», en Pappas (1979), pp. 1-23.

— (1980): «The internalist conception of justification», en French *et al.* (1980), pp. 27-51.

GOODMAN, N. (1952): «Sense and certainty», *Philosophical Review*, 61, 160-167. Reimpreso en Chisholm y Swartz (1973).

— (1973): *Fact, Fiction and Forecast*, Bobbs-Merrill, Indianapolis, Ind.

GORDON, D. (1984): «Knowledge, reliable methods and Nozick», *Analysis*, 44, 30-33.

GRANDY, R. (1973): «Reference, meaning and belief», *Journal of Philosophy*, 70, 439-452.

GRICE, P. (1961): «The causal theory of perception», *Proceedings of the Aristotelian Society*, sup. vol. 35, 121-152. Reimpreso en Swartz (1965).

GUTTENPLAN, S. (ed.) (1975): *Mind and Language*, Clarendon Press, Oxford.

HABERMAS, J. (1972): *Knowledge and Human Interests*, traducción de J. J. Shapiro, Heinemann, Londres. [Edición en castellano: *Conocimiento e interés*, traducción de M. Francisco-Jiménez, 3.ª ed., Taurus, Madrid, 1988.]

HACKER, P. M. S. (1972): *Insight and Illusion*, Oxford University Press, Oxford.

HAMLYN, D. W. (1970): *The Theory of Knowledge*, Macmillan, Londres.

HARE, R. M. (1963): *Freedom and Reason*, Oxford University Press, Oxford.

HARMAN, G. (1970): «Induction», en Swain (1970), pp. 83-99.

HARRE, R., y MADDEN, E. H. (1975): *Causal Powers*, Blackwell, Oxford.

HEGEL, G. W. F. (1931): *The Phenomenology of Mind*, traducción de J. B. Baillie, Allen and Unwin, Londres; 1.ª ed., 1807.

HEMPEL, C. G. (1942): «The function of general laws in history», *Journal of Philosophy*, 39, 35-48. Reimpreso en Feigl y Sellars (1949).

HOOKWAY, C. (1978): «Indeterminacy and interpretation», en C. Hookway y P. Petit

(eds.), *Action and Interpretation*, Cambridge University Press, Cambridge, pp. 17-41.

HUME, D. (1955): *An Inquiry Concerning Human Understanding*, edición de C. W. Hendel, Bobbs-Merrill (Library of Liberal Arts), Nueva York; 1.ª ed., 1748. [Edición en castellano: *Investigación sobre el conocimiento humano*, traducción de J. Salas Oñueta, 6.ª ed., Alianza, Madrid, 1988.]

—(1967): *A Treatise of Human Nature*, edición de L. A. Selby-Bigge, Clarendon Press, Oxford, 1.ª ed., 1739. [Edición en castellano: *Tratado de la naturaleza humana*, traducción de F. Duque, 2.ª ed., Tecnos, Madrid, 1992.]

HYLTON, P. (1982): «Analyticity and the indeterminacy of translation», *Synthese*, 52, 167-184.

JACKSON, F. (1977): *Perception*, Cambridge University Press, Cambridge. KANT, I. (1961): *Critique of Pure Reason*, traducción de N. Kemp Smith, Mcmillan, Londres; 1.ª ed., 1781. [Edición en castellano: *Crítica de la razón pura*, traducción de P. Ribas, 7.ª ed., Alfaguara, Madrid, 1989.]

—(1967): *Philosophical Correspondence 1759-99*, edición y traducción de A. Zweig, University of Chicago Press, Chicago, III.

KRIPKE, S. (1971): «Identity and necessity», en M. Munitz (ed.), *Identity and Individuation*, New York University Press, Nueva York; pp. 135-64.

—(1981): «Wittgenstein on rules and private language», en I. Block (ed.), *Perspectives on the Philosophy of Wittgenstein*, Blackwell, Oxford, pp. 238-312.

—(1982): *Wittgenstein: On Rules and Private Language*, Blackwell, Oxford.

KVANVIG, J. L. (1984): «Subjective justification», *Mind*, 93, 71-84.

LEHRER, K. (1974): *Knowledge*, Clarendon Press, Oxford.

LEHRER, K., y PAXSON, T. (1969): «Knowledge: undefeated justified true belief», *Journal of Philosophy*, 66, 225-237. Reimpreso en Pappas y Swain (1978).

LEVI, I. (1980): *The Enterprise of Knowledge*, MIT Press, Cambridge, Mass.

LEWIS, C. I. (1946): *An Analysis of Knowledge and Valuation*, Open Court, La Salle, III. El cap. 7 está reimpreso en Chisholm y Swartz (1973); el cap. 8 está reimpreso enSwartz(1965).

—(1952): «The given element in empirical knowledge», *Philosophical Review*, 61, 168-175. Reimpreso en Chisholm y Swartz (1973).

LEWIS, D. (1973): *Counterfactuals*, Blackwell, Oxford.

—(1974): «Radical interpretation», *Synthese*, 23, 331-344.

LOCKE, D.. (1971): *Memory*, Macmillan, Londres.

LOCKE, J. (1961): *An Essay Concerning Human Understanding*, edición de J. Yolton, Everyman's Library, Dent/Dutton; Londres, 1.ª ed., 1690. [Edición en castellano: *Ensayo sobre el entendimiento humano*, traducción de L. Rodríguez Aranda, 2.ª ed., Águilar, Madrid, 1961.]

MCCARTHY, T. (1978): *The Critical Theory of Jürgen Habermas*, MIT Press., Cambridge, Mass. [Edición en castellano: *La teoría crítica de Jürgen Habermas*, traducción de M. Jiménez Redondo, Tecnos, Madrid, 1987.]

MCDOWELL, J. (1982): «Criteria, defeasibility and knowledge», *Proceedings of the British Academy*, 68, pp. 455-479, Oxford University Press, Oxford, 1983.

MACKIE, J. L. (1976): *Problems from Locke*, Clarendon Press, Oxford.

MALCOLM, N. (1958): «Knowledge of other minds», *Journal of Philosophy*, 55, 969-978. Reimpreso en su *Knowledge and Certainty*, Prentice-Hall, Englewood Cliffs, 1963, y en Pitcher (1966).

MARTIN, C. B., y DEUTSCHER, M. (1966): «Remembering», *Philosophical Review*, 75, 161 -196. Reimpreso en Chisholm y Swartz (1973).

MILL, J. S. (1867): *An Examination of Sir William Hamilton's Philosophy*, 3.ª ed., Longmans, Londres.

NIDDITCH, P.H. (ed.) (1968): *The Philosophy of Science*, Oxford University Press, Oxford.

NOZICK, R. (1981): *Philosophical Explanations*, Oxford University Press, Oxford.

OTTMANN, H. (1982): «Cognitive interests and self-reflection», J. B. Thompson y D. Held (eds.), *Habermas: Critical Debates*, Macmillan, Londres, pp. 79-97.

PAPPAS, G. (ed.) (1979): *Justification and Knowledge*, Reidel, Dordrecht.

PAPPAS, G., y SWAIN, M. (eds.) (1978): *Essays on Knowledge and Justification*, Cornell University Press, Ithaca, NY.

PEACOCKE, C. (1983): *Sense and Content*, Clarendon Press, Oxford.

PEARS, D. (1966): «Incompatibilities of colours», en A. G. N. Flew (ed.), *Logic and Language*, 2nd series, Blackwell, Oxford, pp. 112-122.

PHILLIPS GRIFFITHS, A. (ed.) (1967): *Knowledge and Belief*, Oxford University Press, Oxford.

PITCHER, G. (ed.) (1966) *Wittgenstein*, Macmillan, Londres.

PLATÓN (1973): *Theaetetus*, edición y traducción de J. McDowell, Clarendon Press, Oxford. [Edición en castellano: *Teeteto*, en *Platón: Obras completas*, traducción de J. A. Míguez, 2.ª ed., Aguilar, Madrid, 1977.]

POLLOCK, J. L. (1979): «A plethora of epistemological theories», en Pappas (1979), pp. 93-113.

PRICHARD, H. A. (1967): «Knowing and believing», extractos de su *Knowledge and Perception*, Clarendon Press, Oxford, 1950; reimpreso en Phillips Griffiths (1967).

PUTNAM, H. (1975): *Mind, Language and Reality; Philosophical Papers*, vol. 2, Cambridge University Press, Cambridge.

— (1977): «Two dogmas revisited», en G. Ryle (éd.), *Contemporary Aspects of Philosophy*, Oriel Press, Londres, pp. 202-213.

QUINE, W. V. O. (1953): *From a Logical Point of View*, Harvard University Press, Cambridge, Mass. [Edición en castellano: *Desde un punto de vista lógico*, traducción de M. Sacristán, 2.ª ed., Orbis, Barcelona, 1985.]

— (1960): *Word and Object*, MIT Press, Cambridge, Mass. [Edición en castellano: *Palabra y objeto*, Ariel, Barcelona, 1962.]

— (1969): *Ontological Relativity*, Colombia University Press, Nueva York. [Edición en castellano: *La relatividad antológica y otros ensayos*, traducción de M. Garrido y J. Ll. Blasco, Tecnos, Madrid, 1974.]

— (1970): «On the reasons for indeterminacy of translation», *Journal of Philosophy*, 67, 178-183.

— (1973): *The Roots of Reference*, Open Court, LaSalle, 111.

— (1975a): «The nature of natural knowledge», en Guttenplan (1975), pp. 67-81. [Edición en castellano: *Las raíces de la referencia*, traducción de M. Sacristán, Alianza, Madrid, 1988.]

— (1975b): «Reply to Chomsky», en D. Davidson y J. Hintikka (eds.), *Words and Objections*, Reidel, Dordrecht, pp. 302-311.

— (1981): «Reply to Stroud», en P. French, T. E. Uehling y H. K. Wettstein (eds.), *Midwest Studies in Philosophy*, vol. 6, *The Foundations of Analytic Philosophy*, University Of Minnesota Press, Minneapolis, Minn., pp. 473-475.

REICHENBACH, H. (1952): «Are phenomenal reports absolutely certain?», *Philosophical Review*, 61, 147-159. Reimpreso en Chisholm y Swartz (1973).

RESCHER, N. (1973): *The Coherence Theory of Truth*, Clarendon Press, Oxford.

RING, M. (1977): «Knowledge: the cessation of belief», *American Philosophical Quarterly*, 14,51-59.

ROCK, I. (1984): *Perception*, Scientific American Library, Nueva York. [Edición en castellano: *La Percepción*, traducción de J. M. García de la Mora, Prensa Científíca, Barcelona, 1985.]

RORTY, R. (1980): *Philosophy and the Mirror of Nature*, Princeton University Press, Princeton, N.J. [Edición en castellano: *La filosofía y el espejo de la naturaleza*, traducción de J. Fernández Zulaica, 2.ª ed., Cátedra, Madrid, 1989.]

RUSSELL, B. (1907): «On the nature of truth», *Proceedings of the Aristotelian Society 1906-7*, 28-49.

— (1921): *The Analysis of Mind*, Allen and Unwin, Londres. [Edición en castellano: *Análisis del espíritu*, Paidós, Buenos Aires, 1962.]

— (1926): *Our Knowledge of the External World*, 2.ª ed., Allen and Unwin, Londres. [Edición en castellano: *Conocimiento del mundo exterior*, Compañía General Fabril, Buenos Aires, 1964.]

— (1959): *The Problems of Philosophy*, Oxford University Press, Oxford. [Edición en castellano: *Los problemas de la filosofía*, traducción de J. Xirau, 9.ª ed., Labor, Barcelona, 1986.]

RYLE, G. (1949): *The Concept of Mind*, Hutchinson, Londres. [Edición en castellano: *El concepto de lo mental*, Paidós, Buenos Aires, 1967.]

SCHLICK, M. (1936): «Meaning and verification», *Philosophical Review*, 45, 339-369. Reimpreso en Feigl y Sellars (1949).

SELLARS, W. F. (1963): *Science, Perception and Reality*, Routledge and Kegan Paul, Londres. El cap. 5 está reimpreso en Chisholm y Swartz (1973).

— (1973): «Givenness and explanatory coherence», *Journal of Philosophy*, 70, 612-624.

— (1979): «More on givenness and explanatory coherence», en Pappas (1979), pp. 169-182.

SHOEMAKER, S. (1963): *Self-knowledge and Self-identity*, Cornell University Press, Ithaca, NY.

SHOPE, R. (1983): *The Analysis of Knowing*, Princeton University Press, Princeton, NJ.

—(1984): «Cognitivie abilities, conditionals and knowledge: a response to Nozick», *Journal of Philosophy*, 81,29-48.

SMALL, K. (1961): «Professor Goodman's puzzle», *Philosophical Review*, 70, 544-552.

SNOWDON, P. (1981): «Perception, vision and causation», *Proceedings of the Aristotelian Society 1980-1*, 81, 175-192.

STRAWSON, P. F. (1952): *introduction to Logical Theory*, Methuen, Londres.

— (1974): *Freedom and Resentment*, Methuen, Londres.

— (1979): «Perception and its objects», en G. F. Macdonald (ed.), *Perception and Identity*, Macmillan, Londres, pp. 41-60.

STROUD, B. (1968): «Transcendental *arguments*», *Journal of Philosophy*, 65, 241-256. Reimpreso en T. Penelhum y J. Macintosh (eds.), *The First Critique*, Wadsworth, Belmont, Calif, pp. 54-69.

— (1969): «Conventionalism and translation», en D. Davidson y J. Hintikka (eds.), *Words and Objections*, Reidel, Dordrecht; pp. 82-96.

— (1977): *Hume*, Routledge and Kegan Paul, Londres.

— (1984): *The Significance of Philosophical Scepticism*, Clarendon Press, Oxford.

SWAIN, M. (ed.) (1970): *Induction, Acceptance and Rational Belief*, Reidel, Dordrecht.

— (1974): «Epistemic defeasibility», *American Philosophical Quarterly*, 11, 15-25. Reimpreso en Pappas y Swain (1978).

— (1981): «Justification and reliable belief», *Philosophical Studies*, 40, 389-407.

SWARTZ, R. (ed.) (1965): *Perceiving, Sensing and Knowing*, University of California Press, Berkeley, Calif.

SWINBURNE, R. G. (ed.) (1974): *The Justification of Induction*, Oxford University Press, Oxford.

TAYLOR, C. (1972): «The opening arguments of the Phenomenology», en A. Maclntyre (ed.), *Hegel: A Collection of Critical Essays*, Anchor Books, Garden City NY, pp. 151-187.

— (1984): «Foucault on freedom and truth», *Political Theory*, 12, 152-183.

URMSON, J. O. (1953): «Some questions concerning validity», *Revue internationale de Philosophie*, 25, 217-229. Reimpreso en Swinburne (1974).

VAN INWAGEN, P. (1975): «The incompatibility of free will and determinism», *Philosophical Studies*, 27, 185-199. Reimpreso en Watson (1982).

WATSON, G. (ed.) (1982): *Free Will*, Oxford University Press, Oxford.

WEISK.RANTZ, L. (1980): «Varities of residual experience», *Quarterly Journal of Experimental Psychology*, 32, 365-386.

WITTGENSTEIN, L. (1953): *Philosophical Investigations*, Blackwell, Oxford. [Edición en castellano: *Investigaciones filosóficas*, traducción de A. García Suárez y U. Moulines, Crítica, Barcelona, 1988.]

— (1969a): *On Certainty*, Blackwell, Oxford. [Edición en castellano: *Sobre la certeza*, traducción de J. L. Prades y V. Rega, Gedisa, Barcelona, 1987.]

— (1969b): *The Blue and Brown Books*, Blackwell, Oxford. [Edición en castellano: *Los cuadernos azul v marrón*, traducción de F. Gracia Guillen, Tecnos, Madrid, 1968.]

WOODFIELD, A. (ed.) (1982): *Thought and Object*, Clarendon Press, Oxford.

WOOZLEY, A. D. (1953): «Knowing and not knowing», *Proceedings of the Aristotelian Society, 1952-3*, 53, 151-172. Reimpreso en Phillips Griffiths (1967).

WRIGHT, C. J. G. (1984): «Second thoughts about criteria», *Synthese*, 58, 383-405.

ÍNDICE TEMÁTICO*

ALBRITTON, R.: p. 104.
ALSTON, W. P.: p. 84.
aprehensión/percepción directa/indirecta: 5.1.
—compatible con el error: 10.2.
—compatible con el intervalo temporal: 10.2.
—definición: 10.2.
—el fenomenalismo, una teoría de la aprehensión directa: 10.2.
—la inverosimilitud de la doble aprehensión: 11.2.
—la memoria puede ser aprehensión directa: 10.2.
—objetos directos e indirectos de la memoria: 12.2.
—objetos directos internos (datos sensoriales): 11.2.
—una noción diferente de aprehensión directa: 11.6.
aprehensión doble: 5.1, 10.4, 11.2.
analítico/sintético:
—el rechazo de Quine: 6.3, 14.6, 15.4.
—la justificación analítica de la inducción: 13.3.
—lo sintético/a *priori:* 14.3.
—un análisis coherentista: 14.7.
—y empirismo: 14.2.
—y verdad lógica: 14.6.
antirrealismo:
—como respuesta al escepticismo: 1.4, 6.3, 12.5.
—compatible, pero no exigido por el coherentismo: 9.5.
—comprometido con una teoría incorrecta de la percepción: 11.6.
—en teoría del significado: 6.3.
—introducción: 1.4.
—metafísico/perceptivo: 10.5, 11.6.
—pero es una respuesta incompleta: 9.5.
—tipos de: 9.5.
—y el futuro: 13.2.
—y el pasado: 12.5.
—y fundamentalismo: 7.5.
—y la justificación de la inducción: 13.3.
antirrealismo puro:
—comprometido con el fenomenalismo en teoría de la percepción: 11.6.
—introducción: 9.5.
argumento del error:
—descansa en el principio de universalizabilidad: 1.2.
—el coherentismo no proporciona ninguna respuesta: 9.5.
—internalista: 3.5.
—introducción: 1.2.
—no refutado por Nozick: 3.4.
—refutado fácilmente por el externalismo: 9.3.
—respuestas antirrealistas: 1.4.
—usado contra el realismo indirecto: 11.2.
—y el conocimiento por recuerdo: 12.5.
—y las otras mentes: 5.2.
argumento del intervalo temporal: 10.2, 10.4, 12.2.
argumento por analogía (otras mentes): 5.3.
argumentos trascendentales: 1.3, 12.5.
ARISTÓTELES: 13.2.

* Las referencias a temas o autores que aparecen en el texto principal se dan mencionando la sección del texto (v. gr., 1.4); cuando aparecen en las «Lecturas complementarias», se usa el número de página (v. gr., p. 67).

ARMSTRONG, D. M.: 2.3, 9.3, 11.3, pp. 84, 183-184.
asimetrías:
—el coherentismo no contiene tales asimetrías: 8.1.
—genética/continua: 8.5.
—incompatibles con el holismo: 7.3.
—la asimetría del empirismo es compatible con el coherentismo: 8.4, 8.5.
—semánticas y epistemológicas en Quine: 7.2.
—y fundamentalismo: 7.2.
atomismo: véase holismo.
AUSTIN, J. L.: p. 84.
AYER, A. J.: 4.2, 5.6, 6.2-6.3, 11.2, 12.3, 14.2, pp. 84, 116-7, 183-4.

BAKER, G.: 5.6, 5.7, 5.8, pp. 104.
BARKER, S. F., y ACHISTEIN, P.: p. 241.
BERKELEY, G.: 10.3, 10.6, p. 164.
BERLIN, I.: pp. 116-7.
BERNSTEIN, R.: p. 274.
BLACK, M.: 13.3, p. 241.
BLANSHARD, B.: 8.1, 8.2, 9.5, 13.5, 14.7, pp. 149, 164.
BONJOUR, L.: 9.4, 9.5, pp. 164.
BRADLEY, F. J.: 8.1, 8.2, 8.5, 9.5, 14.7, pp. 131, 149,257.
BRANDT, R.: p. 224.
BROAD, C. D.: 14.4-14.5.

CAHN, S. M.: 241.
CARNAP, R.: 6.3.
caso propio (el), como punto de arranque: 4.1, 5.1-5.3.
—el argumento del lenguaje privado (cf.) lo muestra imposible: 5.5-5.7.
—el coherentismo ¿convierte el conocimiento en algo público?: 8.3.
—problemas: 5.4.
causalidad en la memoria: 12.3.
causalidad en la percepción:
—argumentos causales para el realismo indirecto: 10.5, 11.2.
—justificación causal de las creencias perceptivas: 11.5.
—teoría causal de la percepción: 11.4.
—y la imposibilidad de percibir el futuro: 13.1.
causalidad retroactiv: 13.1.
cerebros en cubetas: 1.2, 1.4,3.3.
certeza: 2.1, 3.2, 4.1.
—e infalibilidad: 4.1, 4.3.
cognitivismo crítico: 15.2.
coherencia (definición): 8.1.
color: 10.3, 10.5, 14.3.
compatibilismo: 13.2.
comprensión:
—el antirrealismo como teoria de la comprensión: 1.3, 1.4.
—escepticismo sobre ella: 1.1, 1.2, 1.4, 5.3,5.4.
—y explicación: 11.1, 13.5.
condicionales subjuntivos: 3.1, 3.3, 6.3, 11.1.
conductismo: 5.4.
conocimiento: cf. escepticismo, argumento del error, justificación.
—conocimiento por recuerdo: 12.5, 12.6.
—definición tripartita: 2.1.
—¿implica creencia?: 2.1, 2.2, 9.3.
—¿por qué son conocimiento las creencias perceptivas?: 11.5.
—teoría condicional: 3.1, 3.2.
—y escepticismo: 1.1-1.5.
conocimiento *a priori:* 14.1-14.7.
—de los principios de inferencia: 12.3, 14.1.
—el rechazo de Quine: 14.6.
—lo sintético *a priori:* 14.3.
—¿sólo de verdades necesarias?: 14.5.
—¿sólo de verdades universales?: 14.4.
—y el coherentismo: 14.7.
conocimiento matemático: 14.3.
conocimiento no inferencial: cf. justificación inferencial.
contenidos de aprehensión:
—como modos de ser consciente: 11.3.
—como objetos directos: 10.4.
contradicción: cf. Ley de no contradicción.
CORNMAN, J. W.: 11.3, pp. 149, 183-4, 209.
COUSINS, M., y HUSSAIN, A.: p. 274.
creencia:
—el recuerdo como creencia: 12.3.
—interacción entre creencia y significado: 7.4.
—ninguna creencia es infalible: 4.2.

—percepción como forma de creencia: 11.3.
—y conocimiento: 2.1, 2.4.
—y la experiencia sensorial: 4.2, 8.4.
—y la sensación en la percepción: 11.3.
creencia perceptiva: cf. creencia.
—¿asunto de grado?: 13.1.
—su justificación: 11.5.
—su seguridad antecedente: 8.4, 8.5.
creencia sensorial: 8.4, 8.5.
creencias básicas/no básicas: 4.1-4.3; cf. justificación inferencial.
criterio: 5.4, 5.8.
—de traducción: 7.4.
—el problema del criterio: 15.1-15.5, especialmente 15.2.
CRUSOE, R.: 5.5.
cualidades primarias y secundarias: 10.3, 10.5.

CHISHOLM, R.: 4.2, 11.3, 15.2; p. 84, pp. 116-17, p. 274.
CHISHOLM, R., y SWARTZ, R.: p. 84.

DANCY, J.: pp. 66-7, 149, 274.
datos sensoriales: 11.2, 11.3.
DAVIDSON, D.: p. 131.
de re (creencia): 9.3.
definición ostensiva: 5.6.
definición tripartita del conocimiento: cf. conocimiento, justificación.
—alternativas: 9.3.
—ampliaciones de la definición: 2.3.
—estrategias de respuesta a Gettier: 2.2, 2.3.
—introducción: 2.1.
—los contraejemplos de Gettier: 2.2.
—teoría condicional: 3.1, 3.2.
DENNETT, D.: p. 104.
DESCARTES, R.: 1.2, 1.5, 3.3, 4.3, 9.3.
determinisno: 13.2.
dolor: 5.3-5.7, 15.2.
DRESTK.E, F.: 2.3, 3.2, p. 53.
DUHEM, P.: 6.3, 15.3.
DUMMETT, M. A. E.: 1.4, p. 37, p. 131, p. 164, p. 224, p. 241.
EDWARDS, P.: 13.3, p. 241.
empirismo:
—compatible con el coherentismo puro: 8.4,8.5.
—fundamentalismo clásico, una variante de empirismo: 4.1.
—¿implica fundamental ismo?: 7.2.
—la justificación de nuestra confianza en la creencia perceptual: 11.5.
—variantes más débiles: 4.3.
—y lo *a priori* : 14.2, 14.4, 14.5.
—y lo analítico: 14.3.
—y lo necesario: 14.5-14.7.
empirismo lógico: cf. positivismo lógico
epistemología naturalizada: 15.4.
—no da respuesta al argumento del error: 15.4.
error verbal: 4.2.
escepticismo:
—argumento del error (cf.)
—cerebros en cubetas (cf.)
—el coherentismo no tiene respuesta: 9.5.
—el fundamentalismo como expresión de escepticismo: 5.1, 5.7.
—forma débil de escepticismo: 5.7.
—grados de: 1.1.
—Hume y la inducción: (cf).
—la epistemología naturalizada no tiene respuesta: 15.4.
—la refutación de Nozick: 3.3-3.4.
—la respuesta de Hegel: 15.1.
—respuestas trascendentales: 1.3-1.5, 12.5
—sobre la posibilidad de la epistemología: 15.1-15.5.
—sobre las otras mentes: 5.2-5.4.
—sobre el conocimiento por recuerdo: 12.5.
—una respuesta sugerida: 15.5.
—y el antirrealismo: 1.4, 6.3, 9.5.
—y el fenomenalismo: 6.3.
—y el internalismo: 3.5, 9.2-9.5.
—y el realismo indirecto: 11.2.
—y el realismo inferencial: 11.2.
estados sensoriales e infalibilidad: 4.1-4.2; cf. lenguaje privado, usos comparativos, experiencia inmediata.
ética y epistemología:
—relaciones entre ellas: 1.2, 9.3, 14.4, 15.5.
evidente de un modo directo: 4.2.
EWING, A.C.: 8.1, 8.3, 9.5, 13.5, p. 241.
experiencia inmediata: cf. aprehensión,

creencia, lenguaje privado, externalismo en filosofía de la mente.
—como forma de creencia: 8.4, 11.3.
—como objeto de aprehensión: 10.4, 11.2.
—como un modo de ser consciente: 11.3.
—no tiene conocimiento infalible: 4.2.
—su conocimiento infalible: 4.1.
experimentador caprichoso: 11.4.
explicación:
—e implicación: 8.2.
—lo coherente como lo mutuamente explicativo: 8.1.
—y la inducción: 9.1, 13.5.
—y la necesidad: 8.1.
—y las regularidades experimentadas: 11.1.
externalismo (en filosofía de la mente): cf. internalismo.
—introducción: 9.3.
—rechaza el argumento de la ilusión: 11.3-11.5, 15.5.
externalismo (en teoría de la justificación): cf. internalismo.
—argumentos no concluyentes contra el internalismo: 3.5.
—el coherentismo puede escogerlo: 9.4.
—el internalismo, intuitivamente preferible: 9.4.
—¿lo utiliza Quine?: 15.4.
—no es preciso que el rechazo del principio de universalizabilidad sea externalista: 15.5.
—Nozick como externalista: 3.4, 3.5, 11.5.
—rechazo de estrategias external istas: 11.5, 15.4.
—relacionado con el externalismo en filosofía de la mente: 9.3.
—sus ventajas: 9.2.
—y el argumento del error: 9.2.

falibilismo: 2.3, 4.2, 8.3.
falsedad relevante: 2.3.
fatalismo: 13.2.
fenomenalismo: cf. antirrealismo.
eliminativo-reductivo: 10.6, 12.3, 12.4.
—en metafísica: 9.4-9.5.
—en teoría de la memoria: 12.4.
—en teoría de la percepción: 10.2, 10.6.
—en teoría del significado: 6.3.
—rechazado como no explicativo: 11.1.
—y el antirrealismo metafisico: 11.6.
—y escepticismo: 6.3, 7.5, 11.2.
fiabilidad: 2.3, 8.4, 11.5, 12.6, 13.4.
filosofía primera:
—rechazada por Quine: 15.4.
—y escepticismo: 15.2, 15.4.
física y filosofía: 10.1.
—explicación en física: 11.1.
—inferencia a inobservables: 11.2.
—y la distinción cualidades primarias/secundarias: 10.3, 10.5.
—y la pérdida de la filosofía primera: 15.3-15.5.
—y la teoría causal de la percepción: 11.4.
FIRTH, R.: pp. 84, 149.
FLEW, A. G. N.: 13.l, p. 241.
FØLLESDAL, D.: pp. 116-7.
FOUCAULT, M.: p. 274.
fundamentalismo: cf. regreso de justificaciones, justificación inferencial.
—¿apoyado por el empirismo en teoría del significado?: 6.1-6.3.
—argumentos del regreso de justificaciones en su favor: 4.1.
—conducido al solipsismo: 5.4, 5.7.
—criticado: no existe la infalibilidad: 4.2.
—introducción al fundamentalismo clásico: 4.1.
—¿la única forma de empirismo?: 7.2-7.3, 8.5.
—las asimetrías fundamentacionalistas reflejadas en Quine: 7.2.
—una forma débil de escepticismo: 5.1.
—una teoría holista (por ejemplo, el coherentismo) es preferible a sus asimetrías: 7.3-7.5.
—versiones más débiles que la clásica: 4.3, 5.8, 6.1, 7.4, 7.5.
—vulnerable al argumento del lenguaje privado: 5.5-5.7.
y el regreso internalista de justificaciones: 9.1.
—y la justificación del principio de inferencia: 8.3, 14.1.

—y *lo a priori*: 14.2.
futuro: cf. inducción.
—¿conocido de un modo no inductivo?:
—las concepciones antirrealistas contra las realistas: 13.2.

GARRETT, B.: pp. 66-7.
GETTIER, E. L.: 2.1-2.4, 3.1-3.2, 9.3.
GOLDMAN, A. I.: 2.3, 2.4 94, 9.5 p. 53, p. 164.
GOODMAN, N.: 13.4, 13.5, p. 84, p. 241.
GORDON, D.: pp. 66-7.
grados de verdad: 8.2.
GRANDY, R.: 7.4.
GRICE, H. P. 11.4, p. 208.

HABERMAS, J.: p. 274.
HACKER, P. M. S.: 5.6-5.7, p. 104.
HAMLYN, D. W.: 11.4.
HARE, R. M.: 1.2
HARMAN, G.:9.1, 13.5.
HARRE, R., y MADDEN, E. H.: p. 241.
HEGEL, G. W. H.: 15.1, 15.2, 15.4, 15.5, p. 274.
HEISENBERG, W.: 6.3.
HEMPEL, C. G.: 13.5.
holismo:
—el atomismo de Quine al nivel observacional: 7.1, 7.2.
—el coherentismo como teoría bolista: 8.1, 8.3.
—el holismo de Quine: 6.4.
—es preferible el holismo en epistemología: 7.5.
—es preferible una forma de holismo más completa: 7.4.
—introducido en teoría del significado: 6.3.
—las razones de Quine, insuficientes: 7.3.
—nociones holistas de explicación e implicación contra las atomistas: 8.1, 8.2, 13.5.
—y lo analítico/sintético: 6.3, 14.6, 14.7.
HOLMES, S.: 8.2.
HOOK WA Y, C.: p. 131.
HORNSBY, J.: 2.2.
HUME, D.: 1.2, 12.2, 13.2-13.5, p. 37, p. 257.
HYLTON, P.:p- 131.

idealismo:
—eliminativo/reductivo: 10.6.
—en metafísica: 9.5.
—en teoría de la percepción: 10.6, 11.5, 11.6.
—es preferible el fenomenalismo: 10.6.
—no es necesario para el coherentismo: 9.5.
implicación: 8.1, 13.5.
incorregibilidad: 4.3.
indeterminación (principio de): 6.3.
indeterminación del significado: 6.3, 7.1, 7.4, 8.2, 15.4; cf. infra determinación, analítico/sintético.
indeterminación de la teoría: 6.3, 7.1, 7.4, 8.2, 15.4.
indeterminación de la traducción: 7.1; cf. analítico/sintético.
—argumentos a su favor: 7.4.
—previa a la indeterminación del significado: 7.1, 7.4.
indubitabilidad: 5.1.
inducción: cf. explicación, necesidad.
—definida: 13.1.
—el nuevo rompecabezas: 13.4-13.5.
—la justificación coherentista: 8.3, 13.5.
—las dudas de Hume: 1.2.
—respuestas a Hume: 13.3-13.5.
—y el argumento por analogía: 5.3.
—y la inferencia a la mejor explicación: 9.1, 13.5.
inducción intuitiva: 14.4.
infalibilidad: 4.1,4.2.
—y la percepción directa: 10.2.
infradeterminación de la teoría por parte de la evidencia: 6.3, 7.4, 8.2.
internalismo:
—compatible con el coherentismo: 9.4.
—distintos tipos: 9.3.
—el regreso intemalista de justificaciones: 3.5,9.1,9.2.
—introducción: 3.5.
—intuitivamente preferible al externalismo: 3.5, 9.4.
—y el argumento del error: 3.5, 9.2, 15.5.
—y la justificación inferencial: 9.2.

interpretación comunitaria: 5.5.
—discutida: 5.6.
—y el problema de la inducción: 13.4-13.5.

JACKSON, F.: pp. 183-4, 209.
justificación: cf. teoría de la justificación
—como coherencia, fundamentalismo, justificación inferencia!.
—definición del conocimiento en términos de ella: 2.1.
—definida en términos de conocimiento: 2.2, 3.2, 11.5.
—escepticismo respecto a ella: 1.1-1.4.
—y antirrealismo: 1.4.
justificación inferencial/no inferencial: cf. asimetrías.
—el argumento del regreso de justificaciones (cf.): 4.1.
—el empirismo coherentista no contiene tal tipo de asimetría: 8.5.
—esencial para el fundamentalismo: 4.1,4.3.
—y el argumento del lenguaje privado: 5.7-5.8.
—y el coherentismo: 9.1.
—y el empirismo lógico: 6.2.
—y las propiedades epistemológicas: 4.3.
justificación *prima facie:* 8.4.

KANT, L: 8.4,9.5, 14.3-14.4, 14.7, 15.1.
KRIPKE, S.: 5.5-5.6, 14.4, p. 104, p. 257.
KVANVIG. J. L.: p. 164.

LEHRER, K.: 8.1, 8.3.
LEHRER, K., y PAXSON, T.: 3.2.
lenguaje privado (argumento): 5.5-5.7, 11.2, p. 104.
LEVI, I.: 9.1.
LEWIS, C. I.: 4.1, 4.3, p. 84, pp. 116-17, pp. 183-84.
LEWIS, D.: 3.3, pp. 66-7, p. 131.
Ley de No Contradicción: 6.3, 14.4, 14.6-14.7.
Ley de Tercio Excluso: 6.3, 12.5.
LOCKE, D.: p. 224.
LOCKE, J.: 6.1, 10.3, 10.5, 11.2, 14.2, 14.3, pp. 183-4, p. 257.

MACKIE, J. L.:pp. 183-4.
MALCOLM, N.: 5.3.
MARTIN, C. B., y DEUSTCHER, M.: 12.3.
MCCARTHY. T.: p. 274.
MCDOWELL, J.: 11.4, p. 104, p. 209, p. 274.
memoria: 8.3, 12.1-12.6.
—e imaginación: 12.2.
—fáctica/perceptiva: 12.3.
—realismo directo sobre ella: 12.3.
—realismo indirecto sobre ella: 12.2.
—realismo *versus* antirrealismo: 12.5.
—semejanzas entre memoria y percepción: 12.1.
—teoría de la pura creencia: 12.4.
—y fenomenalismo: 12.4.
—y justificación: 12.6.
memoria perceptiva: 12.3.
—y justificación: 12.5.
MILL, J. S.: 5.3, p. 164.
monismo: 8.4-8.5.
MOORE, G. E.: p. 104.
mundo externo: 5.2, 10.2-10.6, 11.1-11.6.
mundos posibles: 3.3, 14.5.

NAPOLEÓN: 2.3.
navaja de Ockham: 10.3.
negación determinada: 15.1.
necesidad: cf. analítico/sintético.
—e inducción: 13.5.
—natural: 13.5.
—Quine, sobre ella: 14.6.
—un punto de vista coherentista: 14.7.
—y el conocimiento empírico/a *priori*: 14.5.
—y explicación: 8.1.
NEURATH, O.: 15.3.
NOZICK, R.: 3.1-3.5, 9.3, 11.5, 12.5, 15.5, p. 37, pp. 66-7.
objeción de la pluralidad: 8.2-8.4, 13.5.
oraciones/enunciados observacionales:
cf. asimetrías.
—su naturaleza: 6.2.
—su relación con lo no observacional: 6.3, 7.4.
—y la determinación del significado en Quine: 7.2.
—y la traducción: 7.1.

otras mentes: 1.4,5.2-5.8, 15.2.
OTTMANN, H.: p, 274.

PARPAS. G., y SWAIN, M.: p. 53.
pasado: cf. memoria.
PEACOCKE, C: p. 209.
PEARS, D.: p. 257.
percepción: 10.1-10.6, 11.1-11.6, 12.1-12.6; cf. aprehensión directa, realismo directo, realismo indirecto, fenomenalismo.
—de la necesidad: 14.5.
—del futuro: 13.1.
—recuerdo como percepción: 12.1-12.3.
—teorías puras y mixtas: 11.3.
PHILLIPS GRIFFITHS, A.: p. 53.
PLATÓN: 9.3, p. 53.
POLLOCK, J: 9.4-9.5.
posibilidad: cf. necesidad,
positivismo lógico:
PRICHARD, H. A.: p. 53.
principio de caridad: 7.4.
principio de cierre (conocimiento): 1.2, 2.2,3.3.
principio de cierre (justificación): 1.2, 2.2,3.3.
principio de humanidad: 7.4.
principio de universalizabilidad: 1.2, 3.4, 12.5, 15.5.
principio de verificación (del significado): 6.2, pp. 116-17.
—tres teorías verificacionistas: 6.3.
—y el empirismo: 7.3.
—y el fundamentalismo: 6.2.
principios de inferencia: 4.2, 8.3, 13.1-13.4
probabilidad:
—absoluta/relativa: 4.1.
—y certeza: 4.1.
—y la inducción: 13.1.
—y la justificación de la creencia perceptiva: 11.5.
propiedades epistemológicas: 4.3, 9.2.
propiedades trascendentes a toda evidencia: 1.2, 1.4,7.1,9.5.
psicología:
—y la epistemología naturalizada: 15.4.
—y la filosofía de la percepción: 10.1.
PUTNAM, H.: pp. 116-17, p. 257.
QUINE, W. V. O.: 6.2, 6.3, 7.1-7.5, 8.2, 8.4, 13.3, 14.4, 14.7, 15.3-15.5.

razones concluyentes: 2.4, 3.2.
realismo directo: cf. argumento de la ilusión, aprehensión directa.
—comparado con el realismo indirecto: 10.2.
—en teoría de la memoria: 12.3.
—¿es preferible la forma científica?: 11.6.
—formas científicas e ingenuas: 10.3.
—incompatible con el realismo antimetafísico: 11.6.
—introducción: 10.2.
—la forma ingenua, preferible: 10.3.
—recomendado como teoría mixta de la percepción: 11.3.
—se adecua al coherentismo: 11.6.
—y el escepticismo: 11.6.
—y la explicación del error perceptivo: 11.3.
—y la teoría adverbial de la sensación: 11.3.
realismo directo ingenuo: cf. realismo directo.
—¿es preferible al realismo directo científico?: 10.3, 11.6.
—introducción: 10.3.
—sus proponentes: pp. 183-4.
realismo directo científico: cf. realismo directo.
—defensores: pp. 183-4.
—introducción: 10.3.
—problemas: 10.3.
—¿es preferieble al realismo directo ingenuo?: 11.6.
realismo indirecto:
—argumentos en su contra: 11.2.
—argumentos en su favor: 10.4.
—comparado con el realismo inferencial: 10.2, 11.2.
—en teoría de la memoria: 12.2-12.3.
—incompatible con el coherentismo: 11.6.
—introducción: 10.2.
—sus formas (científica e indirecta): 10.5.
—una teoría de la aprehensión doble: 11.2.

realismo indirecto ingenuo: cf. realismo indirecto.
—introducción: 10.5.
—rechazo: 10.5.
—sus defensores: pp. 183-4.
realismo indirecto científico: cf. realismo indirecto, argumento de la ilusión.
—defensores: pp. 183-4.
—inmune a las objeciones escépticas: 11.2
—introducción: 10.5.
—no hay noción clara de aprehensión doble: 11.2.
—preferible al realismo indirecto ingenuo: 10.5.
—su explicación del error perceptivo: 11.3.
—una teoría de la doble aprehensión: 11.2.
realismo inferencial: 10.4, 11.2.
realismo metafisico:
—compatible con el coherentismo: 9.5.
—distinto al realismo perceptivo: 10.2.
—incompatible con el antirrealismo perceptivo (fenomenalismo): 11.6.
—*versus* antirrealismo: 1.4.
realismo perceptivo:
—introducción: 10.2.
—una redefinición frente al fenomenalismo: 10.5.
—y el coherentismo: 11.6.
—y el realismo metafisico: 11.6.
refutabilidad: 2.3, 5.4.
regreso de justificaciones (argumento): 4.1 ; cf. regreso internalista de justificaciones, justificación inferencial.
—no efectivo contra el coherentismo: 9.1.
—propiedades formales y epistémicas: 4.3.
—vicioso, no virtuoso: 4.1.
—y Chisholm: 15.2.
—y el empirismo lógico: 6.2.
—y el fundamentalismo no clásico: 4.3.
regreso internalista de justificaciones: 3.5,9.1-9.2.
REICHENBACH, H.: p. 84.
RESCHER. N.: 8.l, 8.3, p. 149.
RING, M.: p. 53.
ROCK, I.: p. 209.
RORTY, R.: p. 274.
RUSSELL, B.: 3.4, 4.2, 8.2, 11.2, 12.2, 12.5, 14.1.
RYLE, G.: p. 104.

SCHLICK, M.: 6.2.
seguir el rastro de la verdad: 3.1-3.2, 11.5, 12.5; cf. teoría condicional del conocimiento.
seguir una regla: 5.5-5.7,6.1, 13.4-13.5.
seguridad (antecedente/subsiguiente): 8.4-8.5, 11.5.
SELLARS, W. F.: 11.3, 15.4, p. 84, p. 149, pp. 183-4, p. 209.
sensación *versus* creencia en la percepción: cf. teorías de la percepción puras y mixtas, creencia.
separación entre la mente y la conducta: 5.3-5.4.
SHOEMAKER, S.: 11.5.
SHOPE, R.: p. 53, pp. 66-7.
sintético: cf. analítico/sintético.
SKINNER, B. F.: p. 84.
SMALL, K.: 13.4-13.5.
SNOWDON, P.: p. 209.
solipsismo: cf. caso propio.
—argumentos escépticos a su favor: 5.3.
—refutado por el argumento del lenguaje privado: 5.5-5.7.
—una forma de antirrealismo: 9.5, 11.6.
—y el fundamentalismo clásico: 5.8.
STRAWSON, P. F.: 11.4, 13.3, pp. 183-4, p. 209.
STROUD, B.: 15.4.
subjetiva/objetiva (justificación): 9.4; cf. internalismo, externalismo, ética y epistemología,
suerte: 2.1,2.3, 3.2, 9.3.
SWAIN, M.: 2.3, p. 53.
SWINBURNE, R. G.: p. 241.

TAYLOR, C.: p. 274.
teoría adverbial de la sensación: 11.3.
teoría causal del conocimiento: 2.4, 3.2.
teoría causal de la justificación: 2.3, 2.4, 3.2,4.3, 15.4
—en teoría de la percepción: 11.5.
teoría condicional del conocimiento: 3.1-3.5.

—introducción: 3.1.
—su respuesta al escepticismo: 3.3-
—tal respuesta es rechazada: 3.4, 15.5.
—una teoría externalista: 3.5.
—y el conocimiento por recuerdo: 12.5.
—y la justificación de la memoria perceptiva: 11.5.
teoría de la justificación como coherencia: 8.3.
—compatible con el realismo: 9.5.
—compatible con idealismo y antirrealismo: 9.5.
—compatible con internalismo y externalismo: 9.4.
—debe preferir el realismo directo científico: 11.6.
—el coherentismo puro es compatible con el empirismo: 8.5.
—es preferible el coherentismo puro: 8.5.
—la justificación de las creencias perceptivas: 11.5.
—mejor adecuada al realismo directo: 11.6.
—no responde al escéptico por sí misma: 9.5.
—teorías puras y débiles: 8.4.
—y la justificación de las creencias de memoria: 12.6.
—y la teoría de la verdad como coherencia: 8.3.
—y lo empírico/a *priori:* 14.7.
—y los principios de inferencia: 8.3.
teoría de la verdad como coherencia:
—compatible con el realismo: 8.5.
—defendida contra la objeción de la pluralidad: 8.3.
—defensa e introducción: 8.1-8.2.
—definicional/criterial: 8.2.
—y el idealismo: 9.5.
—y la teoría de la justificación como coherencia: 8.3.
—y la teoría de la verdad como correspondencia: 8.2.
teoría de la verdad como correspondencia: 8.2.
teorías empiristas del significado: 6.2-6.3.
—compatibles con el holismo: 7.3.
—compatibles con el realismo/antirrealismo: 7.4.
—el verificacionismo no es la única: 7.3.
teorías puras/mixtas de la percepción: 11.3.
testimonio de los demás: 8.3, 11.5.
traducción: 7.1-7.2, 7.4, 11.5; cf. indeterminación de la traducción.
—criterios: 7.4.
traducción radical: 7.1-7.2.
—su indeterminación: 7.4-7.5.

URMSON. J. O.: 13.3.
usos comparativos y no comparativos de términos de apariencia: 4.2.

VANlNWAGEN. P.: p. 241.
verdad contingente:
—definición: 4.2, 14.5.
«verdul»: 13.4-13.5.
verificación fuerte/débil: 6.2.
visión ciega: 11.3.

WEISKRANTZ, L.: p. 209.
WITTGENSTEIN. L.: 2.2, p. 104.
—como fundamentalista: 5.8.
—el lenguaje privado y seguir una regla: 5.5-5.8.
—y el nuevo rompecabezas de la inducción: 13.4, 13.5.
—y los criterios: 5.4.
WOODFIELD, A.:p. 164.
WOOZLEY, A. D.: 2.1.
WRIGHT, C. J. G.: p. 104.